

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 13:14:12 ; Search time 6222 Seconds
(without alignments)
12699.740 Million cell updates/sec

Title: us-09-817-774-30

Perfect score: 4879
Sequence: 1 aactgtcctcctccacaa.....atctaaacaatttaccg 4879

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hlc:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estlun:*
16: em_estlom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrl:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	608	12.5	625	17 B30087	B30087 T25J13TFB T
C 2	441.2	9.0	447	17 BH170465	BH170465 SALK_0029
C 3	399.8	8.2	631	9 AU238585	AU238585 AU238585
C 4	299.4	6.1	835	17 BH8303	BH8303 F24B19-Sp6
C 5	297	6.1	875	17 BH248541	BH248541 BGCAC47TR
C 6	282.2	5.8	399	10 AV789413	AV789413 AV789413

C 7	280.4	5.7	762	17 BH722372	BH722372 BOH060TR
C 8	258.8	5.3	313	10 BE522479	BE522479 M26B4XTM
C 9	258.8	5.3	397	10 AV820534	AV820534 AV820534
C 10	252.8	5.2	388	10 AV804475	AV804475 AV804475
C 11	247.4	5.1	543	17 CNS00MP7	AL080905 Arabidops
C 12	231	4.7	421	10 AV810289	AV810289 AV810289
C 13	230.2	4.7	410	10 AV817908	AV817908 AV817908
C 14	224.8	4.6	457	10 AV544837	AV544837 AV544837
C 15	224	4.6	529	10 AV553366	AV553366 AV553366
C 16	224	4.6	554	10 AV830142	AV830142 AV830142
C 17	217.4	4.5	380	10 AV812548	AV812548 AV812548
C 18	215.8	4.4	399	10 AV801981	AV801981 AV801981
C 19	208.4	4.3	342	10 AV561130	AV561130 AV561130
C 20	192.4	3.9	431	17 BH617242	BH617242 SALK_0361
C 21	186.8	3.8	342	10 AV526218	AV526218 AV526218
C 22	185.4	3.8	460	17 BH170488	BH170488 SALK_0029
C 23	177.8	3.6	724	17 B20222	B20222 T25J13-17 T
C 24	177.4	3.6	453	10 AV786514	AV786514 AV786514
C 25	167	3.4	740	17 BH8763	BH8763 F1P7-Sp6 IG
C 26	157.4	3.2	239	10 BE527244	BE527244 M67J15TM
C 27	145	3.0	258	10 BE530521	BE530521 M7B13STM
C 28	133	2.7	347	10 BE522478	BE522478 M26B4XTM
C 29	130.8	2.7	582	14 H76200	H76200 17905 Lambd
C 30	128	2.6	699	9 A1054619	A1054619 COA00011
C 31	127.8	2.6	390	14 H76508	H76508 18313 Lambd
C 32	127.4	2.6	583	9 A1993836	A1993836 701515125
C 33	125.8	2.6	497	12 BG601019	BG601019 EST505914
C 34	125.8	2.6	508	12 BF053536	BF053536 EST438766
C 35	125.8	2.6	517	12 BG096517	BG096517 EST461036
C 36	125.8	2.6	549	12 BG097155	BG097155 EST461036
C 37	125.8	2.6	599	12 BG597429	BG597429 EST496107
C 38	125.8	2.6	630	12 BG593376	BG593376 EST492054
C 39	125.8	2.6	660	12 BF053408	BF053408 EST438638
C 40	125.8	2.6	696	12 BG592974	BG592974 EST491652
C 41	125.8	2.6	700	12 BG597158	BG597158 EST495835
C 42	125.8	2.6	722	12 BG598860	BG598860 EST503760
C 43	125.8	2.6	742	12 BG599568	BG599568 EST504463
C 44	125.8	2.6	749	12 BG596307	BG596307 EST494985
C 45	125	2.6	686	12 BG135632	BG135632 EST468524

ALIGNMENTS

RESULT 1
LOCUS B30087 625 bp DNA linear GSS 13-Oct-1997
DEFINITION T25J13TFB TAMU Arabidopsis thaliana genomic clone T25J13, DNA
sequence.
ACCESSION B30087
VERSION B30087.1 GI:2516053
KEYWORDS GSS.
SOURCE Thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 625)
Rounsley, S.D., Kelley, J.M., Field, C.E., Craven, M.B., Adams, M.D. and
Venter, J.C.
Use of a BAC End Sequence Database To Identify Minimal Overlaps for
Arabidopsis Genomic Sequencing
Unpublished (1997)
Other GSSs: T25J13TR
Contact: Steve Rounsley
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: rounsley@tigr.org
Seq primer: M13-21
Class: BAC ends

JOURNAL COMMENT

FEATURES High quality sequence stop: 625.
Location/Qualifiers
1. 625
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="T25J13"
/sex="hermaphrodite"
/note="Vector: BelobACII; site_1: HindIII; site_2: HindIII
; Produced by Rod Wing"

BASE COUNT 217 a 125 c 115 g 168 t
ORIGIN

Query Match 12.5%; Score 608; DB 17; Length 625;
Best Local Similarity 98.4%; Pred. No. 2.5e-105;
Matches 614; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 2116 CACATCCATTTTGGCAATCTTTCTCAGTGTGTTTCCGCGTCCCTTGGAGG 2175
|||||
DB 625 CACAAATCCATTTTGGCAATCTTTCTCAGTGTGTTTCCGCGTCCCTTGGAGG 566
|||||
OY 2176 CATGAGTTGACCCCTCGAATGGTAAGAGCTTTGACATCAGGTCTTACTAATTGCA 2235
|||||
DB 565 CATGAGTTGACCCCTCGAATGGTAAGAGCTTTGACATCAGGTCTTACTAATTGCA 506
|||||
OY 2236 ATTCGGAATGATGTTCTGGGAGTCTTGCAGTACAGTACATTAACAGATTTTC 2295
|||||
DB 505 ATTCGGAATGATGTTCTGGGAGTCTTGCAGTACAGTACATTAACAGATTTTC 446
|||||
OY 2236 TGGATTAATCTTAATTAATCTCACTCGAGACATAGTCTGGTCTTCTTAAGTC 2355
|||||
DB 445 TGGATTAATCTTAATTAATCTCACTCGAGACATAGTCTGGTCTTCTTAAGTC 386
|||||
OY 2356 CTATATTCATTAGAGAGCTTCACCTCTTTCGATTGATTCGATCTTCTTAACC 2415
|||||
DB 385 CTATATTCATTAGAGAGCTTCACCTCTTTCGATTGATTCGATCTTCTTAACC 326
|||||
OY 2416 AGTATGAATTAATGCGAAATATCTGATTCATGCTGTGAACACATCCTGATGCTG 2475
|||||
DB 325 AGTATGAATTAATGCGAAATATCTGATTCATGCTGTGAACACATCCTGATGCTG 266
|||||
OY 2476 TGTATGTCACAAATTTCTTGTGTGGAAGCTGTTATTTGGAACACATGCAATTCGAC 2535
|||||
DB 265 TGTATGTCACAAATTTCTTGTGTGGAAGCTGTTATTTGGAACACATGCAATTCGAC 206
|||||
OY 2536 ATGACCGAGATGCTGCTGTGATCATATATAAATTTATATCATCATCTCTCATATAGG 2595
|||||
DB 205 ATGACCGAGATGCTGCTGTGATCATATATAAATTTATATCATCATCTCTCATATAGG 146
|||||
OY 2596 GTACAGAGGAAGAAGCCAGTAGTATTAACAAGATTTCTCTAGCTTTGGCTAATTT 2655
|||||
DB 145 GTACAGAGGAAGAAGCCAGTAGTATTAACAAGATTTCTCTAGCTTTGGCTAATTT 86
|||||
OY 2656 CTTGCCAGATTCCTTTATATGCTAAGGACAGACATATTTATTCAGATTTGTTACTG 2715
|||||
DB 85 CTTGCCAGATTCCTTTATATGCTAAGGACAGACATATTTATTCAGATTTGTTACTG 26
|||||
OY 2716 GCTTAACACATGAACCTTATAGCTT 2739
|||||
DB 25 GCTTAACACATGAACCTTATAGCTT 2

RESULT 2
BH170465/c 447 bp DNA linear GSS 03-OCT-2001
LOCUS BH170465
DEFINITION SALK_002937 Arabidopsis thaliana TDNA insertion lines Arabidopsis
thaliana genomic clone SALK_002937, DNA sequence.
ACCESSION BH170465
VERSION BH170465.1 GI:15905840
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckersalk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of Atlg50430.
Class: TDNA tagged.

FEATURES
source
Location/Qualifiers
1. 447
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_002937"
/clone.lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html

BASE COUNT 159 a 75 c 96 g 117 t
ORIGIN

Query Match 9.0%; Score 441.2; DB 17; Length 447;
Best Local Similarity 99.3%; Pred. No. 1.2e-73;
Matches 443; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3534 CTAGCTTCCTTAAGCTGATGATGAGTATCTCATTTATACAGTTACACATTA 3593
|||||
DB 446 CTAGCTTCCTTAAGCTGATGATGAGTATCTCATTTATACAGTTACACATTA 387
|||||
OY 3594 AAAATGAGCTGTTGATCATCATCAATGCTTTGAGTGTGAGTGGGATGGCTGTAT 3653
|||||
DB 386 AAATGAGCTGTTGATCATCATCAATGCTTTGAGTGTGAGTGGGATGGCTGTAT 327
|||||
OY 3654 TTCCATTATGTTCCGTGAGATCTTAAGTCTTCTTGTGACCGTACCGGCTCTTCGAT 3713
|||||
DB 326 TTCCATTATGTTCCGTGAGATCTTAAGTCTTCTTGTGACCGTACCGGCTCTTCGAT 267
|||||
OY 3714 AACGTAATAATACATCATCTATTTGCTATATTCGAAGCTCTTTCGAAGTAAGCA 3773
|||||
DB 266 AACGTAATAATACATCATCTATTTGCTATATTCGAAGCTCTTTCGAAGTAAGCA 207
|||||
OY 3774 CTTAACAATTAATCATCTTTGTTTAAACGATCTTTGGATCTCTACGATATTTTC 3833
|||||
DB 206 CTTAACAATTAATCATCTTTGTTTAAACGATCTTTGGATCTCTCTACGATATTTTC 147
|||||
OY 3834 TCACCTTCCTTCCTTTGATGAGGACCAAGAGAGAGATGACCATCCGATCAAGTAAG 3893
|||||
DB 146 TCACCTTCCTTCCTTTGATGAGGACCAAGAGAGAGATGACCATCCGATCAAGTAAG 87
|||||
OY 3894 CACCAATCTAAATGATTTTATACATCTCTTTGATCTTATCTGTGATTTGCTGAC 3953
|||||
DB 86 CACCAATCTAAATGATTTTATACATCTCTTTGATCTTATCTGTGATTTGCTGAC 27
|||||
OY 3954 ATATGCGGTTTGTGTTGTAATAA 3979
|||||
DB 26 ATATGCGGTTTGTGTTGTAATAA 1

RESULT 3
 LOCUS AU238585/c 631 bp mRNA linear EST 01-APR-2002
 DEFINITION AU238585 RAF17 Arabidopsis thaliana cDNA clone RAF17-42-H07 5', mRNA sequence.
 ACCESSION AU238585
 VERSION AU238585.1 GI:19877754
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 631)
 Seki, M., Narusaka, M., Ishida, J., Kamiji, A., Setou, M., Nakajima, M., Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y., and Shimozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA
 Unpublished (2002)
 JOURNAL Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rtr.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FUC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified Bluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/Plant/index_e.html) for further details.

FEATURES
 source location/Qualifiers
 1..631
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="RAF17-42-H07"
 /clone_1lb="RAF17"
 /lab_host="DH10B"
 /note="Site 1: BamHI; Site 2: SalI; Subtraction Library. The sequence was obtained from samples subjected to dehydration-treated (1, 2, 5, 10 and 24 hr) and rehydration-treated (1, 2, 5, 10, and 24 hr after dehydration treatment)."
 BASE COUNT 178 a 114 c 134 g 205 t
 ORIGIN

Query Match 8.2%; Score 399.8; DB 9; Length 631;
 Best Local Similarity 99.3%; Pred. No. 8.1e-66;
 Matches 412; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 AACTGTCCTTCTCCACACAGAGATCTCTTCCGGCTGATGATTGGTTAAGG 60
 Db 414 AACTGTCCTTCTCCACACAGAGATCTCTTCCGGCTGATGATTGGTTAAGG 60
 QY 61 ATGTGATGTCGCAATTAATCTGTTCTTTATCATCATCATCATAGGACACTA 120
 Db 355 ATGTGATGTCGCAATTAATCTGTTCTTTATCATCATCATCATAGGACACTA 120
 QY 121 ATTGCATCTCTTAACATCTAATAATTTCTGGAATGAAAAAACTGGAACCCAAA 180
 Db 295 ATTGCATCTCTTAACATCTAATAATTTCTGGAATGAAAAAACTGGAACCCAAA 180
 QY 181 GAATTCACACAGACAGAAAACTAGATCTAGTTTCGCCGGAGTTACCAAAAACAT 240
 Db 235 GAATTCACACAGACAGAAAACTAGATCTAGTTTCGCCGGAGTTACCAAAAACAT 240
 QY 241 AGAAAAATGATGATCTTACGCAATTCCTGGAATTTCTTCCAAAGCAAG 300
 Db 175 AGAAAAATGATGATCTTACGCAATTCCTGGAATTTCTTCCAAAGCAAG 116

QY 301 CCAATGCTTCGCGAGTAAGAAATTTGTAACATATGATCGGATCAAGTGAAGAA 360
 Db 115 CCAATGCTTCGCGAGTAAGAAATTTGTAACATATGATCGGATCAAGTGAAGAA 360
 QY 361 TTGATTTTGTATGTCGATTCATCATCATGTTGTTGAATACAGATGAGAG 415
 Db 55 TTGATTTTGTATGTCGATTCATCATCATGTTGTTGAATACAGATGAGAG 415

RESULT 4
 LOCUS B18303 835 bp DNA linear GSS 09-AUG-1997
 DEFINITION F24B19-Sp6 IGF Arabidopsis thaliana genomic clone F24B19, DNA sequence.
 ACCESSION B18303
 VERSION B18303.1 GI:2316207
 KEYWORDS GSS.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 835)
 Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and Ecker, J.
 BAC End Sequences at ATGC
 Unpublished (1997)
 JOURNAL Contact: Ecker J.
 Arabidopsis Thaliana Genome Center
 University of Pennsylvania
 Dept. of Biology, University of Pennsylvania, Philadelphia, PA 19104
 Tel: 215-898-9384
 Fax: 215-898-8780
 Email: jecker@atgenome.bio.upenn.edu
 Seg primer: Sp6
 Class: BAC ends
 High quality sequence start: 104
 High quality sequence stop: 116.
 Location/Qualifiers
 1..835
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="F24B19"
 /clone_1lb="IGF"
 /sex="hermaphrodite"
 /note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI. Produced by Thomas Altman"

BASE COUNT 187 a 183 c 163 g 258 t 44 others
 ORIGIN

Query Match 6.1%; Score 299.4; DB 17; Length 835;
 Best Local Similarity 74.6%; Pred. No. 8.6e-47;
 Matches 515; Conservative 0; Mismatches 152; Indels 23; Gaps 12;

QY 3280 GCAGAAATCTGTCGATTCATTAACATATGATGATGATGATGATGATGATGAT 3339
 Db 118 GCTGAAATCTGTCGATTCATTAACATATGATGATGATGATGATGATGATGAT 3339
 QY 3340 AGGACAAACGGGAATGTTGTTGGGAAGAGCCCGCTCAAGGTGATCAATAA 3399
 Db 178 ATGACATATCTGGAATAGTTGGGCTTGGGANG-GGCGCTCAAGGTGATCAATAA 236
 QY 3400 GCCCAACTAGCTTATGAACTTATGATATTCCTTTTCGACCTAAACATTTGATG 3459
 Db 237 GCTCAACTAGCTTATGAACTTATGATATTCCTTTTCGACCTAAACATTTGATG 289
 QY 3460 TATGAATCTCTCTTGGTCTCGGTTGTTGGAGATGTTGGGCTGTATACAAATC 3519
 Db 290 ATGTATGTCGTCTCTGTCCTCCTGCGTTGTCGATGTTGGGCTG-TTACTACA---C 344
 QY 3520 TGGTGAACCTAAACATGATCTCTCTTAACGCTGATGATGATGATGATGATGAT 3579

Db 345 TCGTGTGATACAGTACGCT--CTCTNAGCTGTGAT--GTGAGACTA--TCTACATTA-- 399
 3580 CAGTTACACCAATAAATAGAGTGTTCATCATCATAGGAGCTTTGAATGTCAGGTGG 3639
 Db 400 CAGTTACACCAATAAATAGAGTGTTCATCATCATAGGAGCTTTGAATGTCAGGTGG 458
 3640 GATTGCTGCTCATTTTCATTATGTTCCGAGATCTTAAAGTCTTTCTTCTGACCGTAC 3699
 Db 459 GGATGGCTGCTCATTTTCATTATGTTCCGAGATCTTAAAGTCTTTCTTCTGACCGTAC 517
 3700 CGGCTCTCTGATACAGTAAATFACTATCATCTATTTCTTATTCGAGCTCTTCTT 3759
 Db 518 CTGCTCTCTGATACAGTAAATFACTATCATCTATTTCTTATTCGAGCTCTTCTT 577
 3760 GCAAGTAAAGCAGCTTAAATATATATATATATATATATATATATATATATATATAT 3819
 Db 578 GCAAGTAAAGCAGCTTAAATATATATATATATATATATATATATATATATATATAT 637
 3820 CTAGCTATATTTCTGACCCCTTCTTCTTCTGATGAGCCAG--AGAGACGATGACCGAT 3878
 Db 638 CTAGCTATATTTCTGACCCCTTCTTCTTCTGATGAGCCAG--AGAGACGATGACCGAT 697
 3879 GCCATCAAAATATAGCAACCAATATATATATATATATATATATATATATATATATAT 3938
 Db 698 GCCATCAAAATATAGCAACCAATATATATATATATATATATATATATATATATATAT 755
 3939 TCTGATTTGCTGACAAATATGCGGTTTGT 3968
 Db 756 TGAATGTGACAAATATGCGGTTTGT 785

RESULT 5
 LOCUS BH248541 875 bp DNA linear GSS 26-NOV-2001
 DEFINITION BOGAC47TR BOGA Brassica oleracea genomic clone BOGAC47, DNA
 ACCESSION BH248541
 VERSION BH248541.1 GI:17070041
 KEYWORDS GSS.
 SOURCE Brassica oleracea.
 ORGANISM Brassica oleracea.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 875)
 Town, C.D., Van Aken, S., Utterback, T. and Fraser, C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 Unpublished (2001)
 Other GSSs: BOGAC47TF
 COMMENT Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: Sheared ends.
 FEATURES
 source 1..875
 Location/Qualifiers
 /organism="Brassica oleracea"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone="BOGAC47"
 /clone_1b="BOGA"
 /note="Vector: pBOS1, Site_1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pBOS1 using BstXI linkers"
 BASE COUNT 257 a 144 c 187 g 287 t
 ORIGIN

Query Match 6.1%; Score 297; DB 17; Length 875;
 Best Local Similarity 65.4%; Pred. No. 2.4e-46;

Matches 609; Conservative 0; Mismatches 250; Indels 72; Gaps 9;
 QY 2450 GCTGTGTAACACCATCTGATGCTGTGTATGTACCAAAATCTTCTGGGGAGCTGG 2509
 Db 1 GCTGTGTAACACCATCTGATGCTGTGTATGTACCAAAATCTTCTGGGGAGCTGG 60
 QY 2510 TTATTTGAACACCATCTGATGCTGTGTATGTACCAAAATCTTCTGGGGAGCTGG 2569
 Db 61 TTATTTGAACACCATCTGATGCTGTGTATGTACCAAAATCTTCTGGGGAGCTGG 120
 QY 2570 TTTATATCATCATCTCTCATATAGGTGACAAAGAAAGCAAGTATATATACAAA 2629
 Db 121 ATTACATCAC--ATGACACAAA 141
 QY 2630 GTTCTGACAGTGTGGATATATATATATATATATATATATATATATATATATAT 2689
 Db 142 TTTGTCGACATATTTGGCATATATATATATATATATATATATATATATATATAT 201
 QY 2690 GCAT 2749
 Db 202 GCAATATGTTGATTTTGG--AGCTTTATGATTTGATTTTTCAT 247
 QY 2750 GCTTAT 2809
 Db 248 ACTTAT 307
 QY 2810 TGGCTCTGCTACACTTCTCCAGGAGTACCTTGTGAACACCCCTCGAATCGGAA 2869
 Db 308 TGGCTCTGCTACACTTCTCCAGGAGTACCTTGTGAACACCCCTCGAATCGGAA 367
 QY 2870 CTCGAGTGCATATATTTGAATTTGCTTACAAAGAAATCTCTTGTGTATATCA 2929
 Db 368 CTCGAGTGCATATATTTGAATTTGCTTACAAAGAAATCTCTTGTGTATATCA 427
 QY 2930 ATGAGAACTTTAGATACCACTTTTATACGAAAAAAGCAAGTATGAGCTAA 2989
 Db 428 TACCCTGCTGTTACAAATGAAAAAGCTTAAAGTATACCACCTTCGAGGAAAGCAA 487
 QY 2990 -----ATGTGAAATCCCAATTTCTAGAAATG--CACTTACTCTTGTATGAT 3039
 Db 488 AATCAATTCAGGTTTCTAGAGTATGAGAAATGCTTTTAACTGCTTAAATGAAA 547
 QY 3040 TACTTGTATCTCCATTAACAAAGAGGAGAGAGTAAGTATGATTTGATCAAA 3099
 Db 548 TACTTGTATCTCCATTAACAAAGAGGAGAGTAAGTATGATTTGATCAAA 607
 QY 3100 TGCCTGTTGATATATATGCTATGATTTGATTTGATTTGATTTGATTTGATTTGAT 3159
 Db 608 TCCCTGATATCTAT--GGCTCAAGATTAATGATTTGATTTGATTTGATTTGATTTGAT 665
 QY 3160 AAGAACTGTTCAATGAACACACCTGATGCTATGATTTGATTTGATTTGATTTGAT 3218
 Db 666 AATCTATATATGAT---ACGCTATGATGATGATTAATTTGATTTGATTTGATTTGAT 721
 QY 3219 TCCGTACCTAGACTTCTATATATATATATATATATATATATATATATATATAT 3277
 Db 722 TCTATCTCTGATCTAT 781
 QY 3278 TTGAGGAATCTGTCATTTATATATATATATATATATATATATATATATATATAT 3337
 Db 782 TGGCGGGGATTTCTGATTTATATATATATATATATATATATATATATATATATAT 840
 QY 3338 GAGAGCAACAGGGAATGTTGTTGGGG 3368
 Db 841 AGAAGCAACAGGGAATGCTGGTTGGGG 871
 RESULT 6
 LOCUS AV789413 399 bp mRNA linear EST 29-MAR-2002
 DEFINITION AV789413 RAF16 Arabidopsis thaliana cDNA clone RAF106-83-122 3',
 mRNA sequence.
 ACCESSION AV789413

VERSION AV789413.1 GI:19823396
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 399)
 AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamuya, A., Setou, M., Nakajima, M.,
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.,
 and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 JOURNAL Unpublished (2002)
 COMMENT Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rtc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). This clone is in a
 modified pBluescript vector as a SclI/XhoI insert. Please visit our
 web site (http://www.gsc.riken.go.jp/e/Plant/index_e.html) for
 further details.

FEATURES
 source 1..399
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="RAF06-83-122"
 /clone_id="RAF06"
 /dev_stage="Plants at various developmental stages from
 germination to mature seeds"
 /lab_host="DH10B"
 /note="Site_1: SclI, Site_2: XhoI; subjected to
 dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
 hr) treatments"

BASE COUNT 144 a 69 c 75 g 111 t
 ORIGIN

Query Match 5.8%; Score 282.2; DB 10; Length 399;
 Best Local Similarity 98.7%; Pred. No. 1.8e-43;
 Matches 295; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 3978 AAGTATGGAATATTTGGAAGCTATTTGTGAG-AAGTCAATACAGATCATTCGGG 4036
 |||||||
 DB 299 AAGTATGGAATATTTGGAAGCTATTTGTGAGAAAGTCAATACAGATCATTCGGG 240
 |||||||
 QY 4037 AATTATGATGTAAGAGAGCTGTTCTCTATTTTCTACTTATTACGTTAATTCGAA 4096
 |||||||
 DB 239 AATTATGATGTAAGAGAGCTGTTCTCTATTTTCTACTTATTACGTTAATTCGAA 180
 |||||||
 QY 4097 CGTTGGAATCATCAAAAGACCGTGCACAAACAAATGCAATGATGATGATGATC 4156
 |||||||
 DB 179 CGTTGGAATCATCAAAAGACCGTGCACAAACAAATGCAATGATGATGATGATC 120
 |||||||
 QY 4157 TTTTGCGATTTGATGCTATAGCTTTTCAATCTCTAGCTACGCTTATGATATTCCTTA 4216
 |||||||
 DB 119 TTTTGCGATTTGATGCTATAGCTTTTCAATCTCTAGCTACGCTTATGATATTCCTTA 60
 |||||||
 QY 4217 GATTATCAAGTATAGCTGCGCTTTTCTAATTTATTTATTTATTTATTTGATATTCAG 4275
 |||||||
 DB 59 GATTATCAAGTATAGCTGCGCTTTTCTAATTTATTTATTTATTTGATATTCAG 1

RESULT 7
 BH722372/c 762 bp DNA linear GSS 20-FEB-2002
 LOCUS BOH060TR BO_2_3_KB Brassica oleracea genomic clone BOH060, DNA
 DEFINITION sequence.
 ACCESSION BH722372
 VERSION BH722372.1 GI:19825496

KEYWORDS GSS.
 SOURCE Brassica oleracea.
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Brassica.
 REFERENCE 1 (bases 1 to 762)
 AUTHORS Town, C.D., Van Aken, S., Uterback, T., and Fraser, C.M.
 JOURNAL Whole genome shotgun sequencing of Brassica oleracea
 Unpublished (2001)
 COMMENT other-GSS: BOH060TR
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.

FEATURES
 source 1..762
 /organism="Brassica oleracea"
 /strain="T01000DH3"
 /db_xref="taxon:3712"
 /clone="BOH060"
 /clone_id="BO_2_3_KB"
 /note="Vector: pPOS1; Site_1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pPOS1 using BstXI linkers"

BASE COUNT 289 a 161 c 129 g 183 t
 ORIGIN

Query Match 5.7%; Score 280.4; DB 17; Length 762;
 Best Local Similarity 67.4%; Pred. No. 3.5e-43;
 Matches 520; Conservative 0; Mismatches 206; Indels 46; Gaps 7;

QY 1199 TCTTCAGCTGCTTTCGCTGCTTAAGAGTGAAGGTCATATCTCCAGCCGGAACCG 1258
 |||||||
 DB 762 TCTTCAGCTGCTTTCGCTGCTTAAGAGTGAAGGTCATATCTCCAGCCGGAACCG 703
 |||||||
 QY 1259 ACCAGTTTCAAGATGATGTTCAATTAGTC-----TTGTGTGTGGGAAAG 1301
 |||||||
 DB 702 GCCGGTTTCAAGATGATGTTCAATTAGTCAGATTTGATGATGATGATGATGATG 643
 |||||||
 QY 1302 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1354
 |||||||
 DB 642 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 583
 |||||||
 QY 1355 GCTTACTTTGACATGACATGACATGACATGACATGACATGACATGACATGACATG 1414
 |||||||
 DB 582 GCTTACTTTGACATGACATGACATGACATGACATGACATGACATGACATGACATG 530
 |||||||
 QY 1415 ACTTGTGTTGGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTT 1474
 |||||||
 DB 529 CATTTCTTTTGAATGAATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTTT 477
 |||||||
 QY 1475 ATGTGATTTAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1534
 |||||||
 DB 476 ATGTGATTTAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 417
 |||||||
 QY 1535 GATCATTGCGGTAATTTTGGGATGATGATGATGATGATGATGATGATGATGATGAT 1594
 |||||||
 DB 416 GATCATTGCGGTAATTTTGGGATGATGATGATGATGATGATGATGATGATGATGAT 357
 |||||||
 QY 1595 TTGTACATTAAGTAAAGTGTCTATTGATGATGATGATGATGATGATGATGATGAT 1654
 |||||||
 DB 356 TTGTACATTAAGTAAAGTGTCTATTGATGATGATGATGATGATGATGATGATGAT 297
 |||||||
 QY 1655 ATTGTTGATTAATTTGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1713
 |||||||
 DB 296 TCAGTTTAAATTTGCTTTTATGATGATGATGATGATGATGATGATGATGATGATG 237
 |||||||
 QY 1714 CATGTGACCTTCATCAAGTACATGATGATGATGATGATGATGATGATGATGATGATG 1773

Db	236	CATGTGCACCTTCATCAAGTAGTATTCTGGTTCATGTGTAACTGATTTATCCATTCAT	177
Qy	1774	TGGGTGGTATTTCGCGATGCATGTCATCTCAATCTTTTGAATTTGGCTTTGGGATTAA	1833
Db	176	TGGGTGGTATTTCGTGTACCCAAATTCCTATTCTTCTGTGAGACTAGTGGTTGGGATTAC	117
Qy	1834	TTTGAGAACT-----CTGGTCTTACTGACCATCTCCCATATAATATTTTGGCTCTTTTC	1888
Db	116	TTTAAAGTCTTAACGTTATTGCCAATGTGTAGTCAATATACCAAGATAATATCTTGC--TTT	59
Qy	1889	TCGACATTTGGAGGTCAAGTTTCAGTAAAGTTTCCAAATTTTGTAGACCTGT	1940
Db	58	TAGCACTAGGAGGTGATGACTCTTCGGAATTCACAGATTTAGTAAGTACGT	7

RESULT	8
LOCUS	BE522479/c
DEFINITION	BE522479 313 bp mRNA linear EST 19-MAR-2001 M26B4XtM Arabidopsis developing seed Arabidopsis thaliana cDNA clone M26B4 3', mRNA sequence.
ACCESSION	BE522479
VERSION	BE522479.1 GI:9780457
KEYWORDS	EST.
SOURCE	thale cress,
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops. 1 (bases 1 to 313)
REFERENCE	White,J.A., Todd,J., Newman,T., Focks,N., Gilke,T., Martinez de Irlaruyza,O., Jaworski,J.G., Ohlroge,J. and Benning,C. A new set of Arabidopsis expressed sequence tags from developing seeds. The metabolic pathway from carbohydrates to seed oil plant physiol. 124 (4), 1582-1594 (2000) 20567808
JOURNAL	Contact: Benning, C
MEDLINE	
COMMENT	

Michigan State University
224 Biochemistry, Michigan State University, East Lansing, MI 48824
, USA
Tel: 517 355 1609
Tel: 517 353 9334
Fax: 517 353 9334
Email: denning@msu.edu
Michigan State University DNA Sequencing Facility Arabidopsis
Biological Resource Center, The Ohio State University, 309 Botany &
Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
6142920603 TEL: 6142929371.

FEATURES	SOURCE
Location/Qualifiers	
1. .313	
/organism="Arabidopsis thaliana"	
/strain="Columbia"	
/db_xref="taxon:3702"	
/clone="M2684"	
/clone_1lb="Arabidopsis developing seed"	
/tissue_type="seed"	
/dev_stage="5-13 days after flowering"	
/lab_host="E.coli"	
/note="Organ: Developing seed; Vector: pBluescript SK-	
Site_1: EcoRI; Site_2: XhoI"	
BASE COUNT	
ORIGIN	
105 a	
60 c	
53 g	
95 t	

[illegible]

Db	214	AAATTATTTGATTTGACGAAGTCGTGTGTCATATTTTCAGTAATTAATGAA	155
Oy	4097	CGTTGGAAATCATCAAAAGACCGTGCCAAAACAAAATGCAAAATGATGCGATAAACATTC	415
Db	154	CGTTGGAAATCATCAAAAGACCGTGCCAAAACAAAATGCAAAATGATGCGATAAACATTC	95
Oy	4157	TTTTGCGATTTGTATGCTAATAGTTTTCAAAATCTAGCTACGCTATGATTTCCCTA	4211
Db	94	TTTTGCGATTTGTATGCTAATAGTTTTCAAAATCTAGCTACCTATGATTTCCCTA	35
Oy	4217	GATTATCAAAAGTAGCGTCCGCTTTTCAATTTT	4250
Db	34	GATTATCAAAAGTAGCGTCCGCTTTTCAATTTT	1

RESULT 9	AY820534/c	AY820534	397 bp	mrna	linear	EST 01-APR-2002
LOCUS	AY820534	AY820534	RAFL1	Arabidopsis thaliana	cDNA clone	RAFL1-10-006 3'
DEFINITION	mRNA sequence.					
ACCESSION	AY820534					
VERSION	AY820534.1					
KEYWORDS	EST. S1.1 GI:19862508					
SOURCE	EST. thale. cress.					
ORGANISM	Arabidopsis thaliana					
	viridiplantae.					
	Streptophyta: Embryophyta: Tracheophyta;					

REFERENCE
AUTHORS
 1 (bases 1 to 397)
 Saito, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
 and Shinozaki, K.
TITLE
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL
 Unpublished (2002)
COMMENT
 Contact: Motoaki Seki

Riken Functional Genomics Research Group
Riken Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambda FLIC-1 vector (Carninci et al., submitted for publication) digested with BamI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES	source
Location/Qualifiers	1. .397
/organism="Arabidopsis thaliana"	
/db_xref="taxon:3702"	
/clone="RAFL11-10-006"	
/clone_11b="RAFL11"	
/dev_stage="plants at various developmental stages from germination to mature seeds"	
/lab_host="DH10B"	
/note="Site 1: BamHI; Site2: SalI; subjected to various treatments (dehydration, cold, high salt, ABA, heat and UV light)"	
. Dark-grown plants"	
BASE COUNT	139 a 71 c 79 g 108 t
ORIGIN	

Query Match	5.3%	Score 258.8	DB 10	Length 397
Best Local Similarity	98.9%	Pred. No. 5.1e-39		
Matches 271	Conservative 0	Mismatches 2	Indels 1	Gaps 1
QY 3978	AAGTATGGGAATAATTGGAAAGCTGTATTTGGAG-AAGTCATAACAGATCATTCGGG	4036		
DB 274	AAAGATATGGGAAATATTGGGAAGCTGTATTTGGAAAGTCAATACAGATCATTCGGG	215		
QY 4037	AATTATTGATTTGATACGAACACTGTTGTTCTCAATTTCTACTATTATTCGTTAATTCGAA	4096		

	Query Match	Best Local Similarity	Score	252.8;	DB	10;	Length	388;
	Matches	265;	Conservative	0;	Mismatches	2;	Indels	1;
								Gaps
QY	3978	AAGTATGGGAATATTGGAACTGATATTGGAC	AAGTCAATACAGATATTCGGC	4036				
Db	269	AAAGTATGGGAATATTGGAACTGATATTGGAC	AAATACAGATATTCGGC	210				

241 AGAAAAATCGATACTTACAC 261

REFERENCE AUTHORS TITLE	JOURNAL MEDLINE COMMENT	FEATURES source	BASE COUNT ORIGIN	Query Match Best Local Similarly Matches 224; Conservative	Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;
haile cress. Arabidopsis thaliana	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 529)	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1
CONTACT: Erika Asamizu The First Laboratory for Plant Gene Research Kazusa DNA Research Institute Yana 1532-3, Kisarazu, Chiba 292-0812, Japan Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/location/Qualifiers	1. 529 Location/Qualifiers	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries DNA Res. 7, 175-180 (2000)	128 a 113 c 116 g 172 t	4.6%; Score 224; DB 10; Length 529; 100.0%; Pred. No. 1.9e-32; 0; Mismatches 0; Indels 0; Gaps 0;	1	
Arabidopsis thaliana	Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S. A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries				

100

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 13:37:02 ; Search time 148 seconds
(without alignments)
10109.972 Million cell updates/sec

Title: US-09-817-774-30

Perfect score: 4879

Sequence: 1 aactgtcctcctccacaa.....atctaaacaatttactg 4879

Scoring table:

IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 segs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2_6/ptodata/2/1na/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/1na/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/1na/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/1na/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/1na/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/1na/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	224	4.6	1496	1 US-08-601-435-1	Sequence 1, Appl
2	224	4.6	1496	2 US-08-931-047-1	Sequence 1, Appl
3	224	4.6	1496	4 US-08-783-202-1	Sequence 1, Appl
4	107.6	2.2	1646	4 US-09-443-041A-19	Sequence 19, Appl
5	91.8	1.9	587	4 US-09-443-041A-5	Sequence 5, Appl
6	89.8	1.8	1870	4 US-09-443-041A-17	Sequence 17, Appl
7	89.2	1.8	1695	4 US-09-443-041A-21	Sequence 21, Appl
8	79.2	1.6	527	4 US-09-443-041A-7	Sequence 7, Appl
9	66.6	1.4	7218	1 US-08-232-463-14	Sequence 14, Appl
10	55.4	1.1	7218	1 US-08-232-463-14	Sequence 14, Appl
11	53.6	1.1	260	1 US-08-248-474-71	Sequence 71, Appl
12	53.6	1.1	260	3 US-08-756-849-71	Sequence 71, Appl
13	49.2	1.0	533	4 US-09-443-041A-3	Sequence 3, Appl
14	46.2	0.9	19124	2 US-08-487-826B-13	Sequence 13, Appl
15	45.6	0.9	624	4 US-09-443-041A-1	Sequence 1, Appl
16	45.4	0.9	8920	2 US-08-446-855A-1	Sequence 1, Appl
17	45.4	0.9	8920	4 US-09-150-741-1	Sequence 1, Appl
18	42.8	0.9	658	4 US-08-998-416-595	Sequence 595, App
19	42.4	0.9	168575	4 US-09-426-290-1	Sequence 1, Appl
20	41.8	0.9	1117	4 US-09-247-373B-33	Sequence 33, Appl
21	41.6	0.9	1066	1 US-08-157-101A-4	Sequence 4, Appl
22	41.4	0.8	4766	5 PCT-US93-07261-10	Sequence 10, Appl
23	41.2	0.8	2747	2 US-08-874-347-1	Sequence 1, Appl
24	41.2	0.8	2747	3 US-09-093-522-1	Sequence 1, Appl
25	41.2	0.8	3095	6 5231168-1	Sequence 1, Appl
26	40.6	0.8	6124	4 US-08-213-419B-3	Sequence 3, Appl
27	40.4	0.8	6124	4 US-08-213-419B-3	Sequence 3, Appl

C 28	40	0.8	6243	2 US-09-056-075-1	Sequence 1, Appl
C 29	39.8	0.8	578	4 US-09-602-877A-95	Sequence 95, Appl
C 30	39.6	0.8	5852	1 US-07-867-106-2	Sequence 2, Appl
C 31	39.4	0.8	8585	1 US-08-030-096-3	Sequence 3, Appl
C 32	39.2	0.8	240	1 US-08-628-417-6	Sequence 6, Appl
C 33	39.2	0.8	29793	4 US-09-302-812-38	Sequence 38, Appl
C 34	39.2	0.8	29793	4 US-09-511-477-38	Sequence 38, Appl
C 35	39.2	0.8	29793	4 US-09-511-507-38	Sequence 38, Appl
C 36	38.8	0.8	1191	2 US-08-658-277-1	Sequence 1, Appl
C 37	38.8	0.8	2222	3 US-09-197-380-1	Sequence 1, Appl
C 38	38.8	0.8	20674	4 US-09-641-638-651	Sequence 651, App
C 39	38.6	0.8	1798	4 US-09-797-906-1	Sequence 1, Appl
C 40	38.6	0.8	30549	4 US-09-134-001C-322	Sequence 322, App
C 41	38.4	0.8	2223	1 US-08-257-073-4	Sequence 4, Appl
C 42	38.4	0.8	2435	4 US-09-306-593-1	Sequence 1, Appl
C 43	38.2	0.8	2383	1 US-08-248-466B-13	Sequence 13, Appl
C 44	38.2	0.8	246240	2 US-08-724-394A-20	Sequence 20, Appl
C 45	38.2	0.8	246240	2 US-08-724-394A-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-601-435-1
Sequence 1, Application US/08601435
Patent No. 5739801
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: DNA sequence coding for a protein
TITLE OF INVENTION: of A. thaliana having a delta-5,7 sterol, delta-7
TITLE OF INVENTION: reductase activity, delta-7 protein, production
NUMBER OF SEQUENCES: 31
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
SOFTWARE:
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/601,435
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9501723
FILING DATE: 15-FEB-1995
APPLICATION NUMBER: FR 9506517
FILING DATE: 01-JUN-1995
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1496 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: 76..1365
US-08-601-435-1
Query Match
Best Local Similarity 4.6%; Score 224; DB 1; Length 1496;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1049 ATGTGACAAATGCTTCATGAGATGCTTCATGACGCTTCTTTGGGA 1108
DB 162 ATGTGACAAATGCTTCATGAGATGCTTCATGACGCTTCTTTGGGA 221
QY 1109 GAATGAGCTTCAAGACCTTATCAACATATGCGCAAGACCCACTTGTGATGCTTGAAT 1168

Accession	Sequence	Position
D8	222 GAATGCGATTCAAGACTTTATCAACATATATGCGCAAGACCCACTTTGATGCTTGAAAT	281
QY	1169 TATATTTTGGCTATGAGACATTTAAGCTATCTTTCAGCTGCTTCCCTGGTAAAGAGT	1228
D8	282 TATATTTTGGCTATGAGACATTTAAGCTATCTTTCAGCTGCTTCCCTGGTAAAGAGT	341
QY	1229 TGAGGGTCCATATCTCCAGCCGAAACGACAGTTTCAAGG	1272
D8	342 TGAGGGTCCATATCTCCAGCCGAAACGACAGTTTCAAGG	385

```

1      RESULT 2
2      US-08-931-047-1
3      : Sequence 1, Application US/08931047
4      : Patent No. 5965417
5      : GENERAL INFORMATION:
6      : APPLICANT:
7      : TITLE OF INVENTION: DNA sequence coding for a protein of
8      : TITLE OF INVENTION: A. thaliana having a delta-5,7 sterol,
9      : TITLE OF INVENTION: delta-7 reductase activity, delta7-Red
10     : TITLE OF INVENTION: protein, production process, strains
11     : TITLE OF INVENTION: of transformed yeasts, uses.
12     : NUMBER OF SEQUENCES: 31
13     : COMPUTER READABLE FORM:
14     : MEDIUM TYPE: Floppy disk
15     : COMPUTER: IBM PC compatible
16     : OPERATING SYSTEM: PC-DOS/MS-DOS
17     : SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
18     : CURRENT APPLICATION DATA:
19     : APPLICATION NUMBER: US/08/931,047
20     : FILING DATE:
21     : CLASSIFICATION: 435
22     : PRIOR APPLICATION DATA:
23     : APPLICATION NUMBER: FR 9501723
24     : FILING DATE: 15-FEB-1995
25     : PRIOR APPLICATION DATA:
26     : APPLICATION NUMBER: FR 9506517
27     : FILING DATE: 01-JUN-1995
28     : INFORMATION FOR SEQ ID NO: 1:
29     : SEQUENCE CHARACTERISTICS:
30     : LENGTH: 1496 base pairs
31     : TYPE: nucleic acid
32     : STRANDEDNESS: single
33     : TOPOLOGY: linear
34     : MOLECULE TYPE: cDNA
35     : ORIGINAL SOURCE:
36     : ORGANISM: Arabidopsis thaliana
37     : FEATURE:
38     : NAME/KEY: CDS
39     : LOCATION: 76..1365
40     : US-08-931-047-1

```

Query Match	4.6%	Score 224	DB 2	Length 1496
Similarity	100.0%	Pred No. 5	1e-47	
Best Local	0	Mismatches	0	Gaps 0
Matches 224	Conservative			

QY	1049	ATGATACAAAGGTTATCAAGAGAGGTCGTTACTCAACACCTTGGCTTTTGGGA	1108
QY	162	ATGGATACAAAGGTTCAACGAGAGGTCCTGTACTCAACACCTTGGCTTTGGGA	221
QY	1109	GATGAGGTTCAAGGACTTATCAACATATGCGCAAGACCACCTTGATGCTTGAAGAAAT	1168
Db	222	GATGAGGTTCAAGGACTTATCAACATATGCGCAAGACCACCTTGATGCTTGAAGAAAT	281
QY	1169	TATATTTTGCATGAGACATTTGAAGCTATTTCCACCTCTCTCGCGGTAAGAAAGT	1228
Db	282	TATATTTTGCATGAGACATTTGAAGGCTATTTCCACCTCTCTCTCGCGGTAAGAAAGT	341
QY	1229	TGAGGGTCCAAATATCCACGACGCGAAGAACGACACAGTTTACAAGG	1272
Db	342	TGAGGGTCCAAATATCTCCACGCGCGAAGAACGACACAGTTTACAAGG	385

```

1      RESULT 3
2      US-08-783-202-1
3      ; Sequence 1, Application US/08783202
4      ; Patent No. 5989881
5      ; GENERAL INFORMATION:
6      ; APPLICANT:
7      ; TITLE OF INVENTION: DNA sequence coding for a protein of A.
8      ; TITLE OF INVENTION: thaliana having a delta-5,7 sterol,delta-7 reductase
9      ; TITLE OF INVENTION: activity, delta7-Red protein, production process, strains
10     ; TITLE OF INVENTION: of transformed yeasts, uses.
11     ; NUMBER OF SEQUENCES: 31
12     ; COMPUTER READABLE FORM:
13     ; MEDIUM TYPE: Floppy disk
14     ; COMPUTER: IBM PC compatible
15     ; OPERATING SYSTEM: PC-DOS/MS-DOS
16     ; SOFTWARE: Patentln Release #1.0, Version #1.30 (EPO)
17     ; CURRENT APPLICATION DATA:
18     ; APPLICATION NUMBER: US/08/783,202
19     ; FILING DATE:
20     ; CLASSIFICATION: 435
21     ; PRIOR APPLICATION DATA:
22     ; APPLICATION NUMBER: FR 9501723
23     ; FILING DATE: 15-FEB-1995
24     ; PRIOR APPLICATION DATA:
25     ; APPLICATION NUMBER: FR 9506517
26     ; FILING DATE: 01-JUN-1995
27     ; INFORMATION FOR SEQ ID NO: 1:
28     ; SEQUENCE CHARACTERISTICS:
29     ; LENGTH: 1496 base pairs
30     ; TYPE: nucleic acid
31     ; STRANDEDNESS: single
32     ; TOPOLOGY: linear
33     ; MOLECULE TYPE: cDNA
34     ; ORIGINAL SOURCE:
35     ; ORGANISM: Arabidopsis thaliana
36     ; FEATURE:
37     ; NAME/KEY: CDS
38     ; LOCATION: 76..1365
39     ;
40     ; US-08-783-202-1

```

Query Match	4.6%;	Score 224;	DB 2;	Length 1496;
Best Local Similarity	100.0%;	Pred. No. 5.1e-47;		
Matches 224;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1049	ATGCTACACAAAGTCTTACAGAAATGCTGCTTACTACAACCTTGGCTCTTTGGGA	1108
Db	162	ATGGTACACAAAGTCTTACAGAAATGCTCTTACTACAACCTTGGCTCTTTGGGA	221
QY	1109	GATGAGATTCAGGACCTTATCAACATATGCGCCAAAGCCACCTTGATTGCTTGGAAAT	1166
Db	222	GATGAGATTCAGGACCTTATCAACATATGCGCCAAAGCCACCTTGATTGCTTGGAAAT	281
QY	1169	TATATTTTTCGTATGAGACATTTTGAAGCTATTTCTCAGCTGCTTCTGCTGTGTAAGAAAGAT	1228
Db	282	TATATTTTTCGTATGAGACATTTTGAAGCTATTTCTCAGCTGCTTCTGCTGTGTAAGAAAGAT	341
QY	1229	TGAGGCTCCAAATATCCAGCGCGGAAATCCGACAGCTTTTAAACAAG	1272
Db	342	TGAGGCTCCAAATATCTCCAGCGCGGAAATCCGACAGCTTTTAAACAAG	385

```

; FILE REFERENCE:
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; APPLICANT: Shen, Jennie
; APPLICANT: Rafalski, Antoni
; APPLICANT: Orozco, Buddy
; APPLICANT: Farnodu, Omolayo O.
; GENERAL INFORMATION:
; Patent No. 6465717
; Sequence 19, Application US/09443041A
US-09-443-041A-19
RESULT 4

```


;; CURRENT APPLICATION NUMBER: US/09/443,041A
;; CURRENT FILING DATE: 1999-11-18
;; PRIOR APPLICATION NUMBER: 60/109,283
;; PRIOR FILING DATE: 1998-11-20
;; NUMBER OF SEQ ID NOS: 33
;; SOFTWARE: Microsoft Office 97
;; SEQ ID NO 19
;; LENGTH: 1646
;; TYPE: DNA
;; ORGANISM: Glycine max
US-09-443-041A-19

Query Match 2.2%; Score 107.6; DB 4; Length 1646;
Best Local Similarity 67.3%; Pred. No. 1.3e-17;
Matches 152; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1049 ATGTACACATATGTTCTCATGAGATGTTCTTCTACTAGACCTTTGGCTTTGGGA 1108
DB 145 ATGTACACATGACTCTTGTCTGAGGATCTGTTCCGACATTCATTTAAGGCA 204
QY 1109 GAATGAGTTCAAGGACTTATCAACATATGAGCAAGCCACCTTGTGTAAGAAAT 1168
DB 205 GAATGAGTTGAGGAGTTGTTGATACATATATGAGCCACCCCTCCACCGCTGCAAAAT 264
QY 1169 TATATTTGCTATGAGGATTTGAAGCTATTTCTTCAAGCTGCTTGCCTGTAAGAGT 1228
DB 265 CATTCAGCTATATGCTGATTTGAGGAGCATCTTCAAGCTTCTTCCCGGAAACCGT 324
QY 1229 TGAGGAGTTCAATATCTCCAGCGGAAACCGACGCTTACAGGTA 1274
DB 325 TTAGGAGCTTATTTCTCCACCGGACGACGCTGTACAGGCA 370

RESULT 5
US-09-443-041A-5
; Sequence 5, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 5
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (282)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (335)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (325)
; OTHER INFORMATION: any nucleotide
; NAME/KEY: unsure
; LOCATION: (530)
; OTHER INFORMATION: any nucleotide
US-09-443-041A-5

Query Match 1.9%; Score 91.8; DB 4; Length 587;

Best Local Similarity 62.4%; Pred. No. 7.9e-14;
Matches 141; Conservative 0; Mismatches 85; Indels 0; Gaps 0;
QY 1049 ATGTACACATATGTTCTCATGAGATGTTCTTCTACTAGACCTTTGGCTTTGGGA 1108
DB 138 ATGTACACATGACTCTTGTCTGAGGATCTGTTCCGACATTCATTTAAGGCA 197
QY 1109 GAATGAGTTCAAGGACTTATCAACATATGAGCAAGCCACCTTGTGTAAGAAAT 1168
DB 198 GAATGAGTTGAGGAGTTGTTGATACATATATGAGCCACCCCTCCACCGCTGCAAAAT 257
QY 1169 TATATTTGCTATGAGGATTTGAAGCTATTTCTTCAAGCTGCTTGCCTGTAAGAGT 1228
DB 258 CATTCAGCTATATGCTGATTTGAGGAGCATCTTCAAGCTTCTTCCCGGAAACCGT 317
QY 1229 TGAGGAGTTCAATATCTCCAGCGGAAACCGACGCTTACAGGTA 1274
DB 318 TTAGGAGCTTATTTCTCCACCGGACGACGCTGTACAGGCA 363

RESULT 6
US-09-443-041A-17
; Sequence 17, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 17
; LENGTH: 1870
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-443-041A-17

Query Match 1.8%; Score 89.8; DB 4; Length 1870;
Best Local Similarity 77.3%; Pred. No. 4.3e-13;
Matches 109; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 2406 TTCTTTAACAGTATGAATAATGCAAGATGATGATTCATGCTGTGAACGATC 2465
DB 750 TGCATAAAGCAATGAAATGATGAGCGGAGTTCAGATTCATGCTGTGAATACGCA 809
QY 2466 CTGATGCTGCTATGATGCAAAATTTCTTGTGGAAGCTGTTATTTGAACCATG 2525
DB 810 TTGATGTTGATCTATGATGCAAAATTTCTTGTGGAAGCTGTTATTTGAACCATG 869
QY 2526 GACATTCAGATGAGGAGT 2546
DB 870 GACATTCAGATGAGGAGT 890

RESULT 7
US-09-443-041A-21
; Sequence 21, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18

[illegible]

[illegible]

```

LOCATION: 1..260
OTHER INFORMATION: /standard name="DB# 317"

```

ADDRESSEE: Knobbe Martens Olson & Bear

```

; STREET: 620 Newport Center Drive 16th Floor
; City: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; US-08-487-826B-13

```

```

Query Match 0.98; Score 46.2; DB 2; Length 19124;
Best Local Similarity 46.7%; Pred. No. 0.13;
Matches 218; Conservative 0; Mismatches 243; Indels 6; Gaps 2;

```

```

QY 1402 AATGTTCCGATGACTTGGTTTGAAGTATTCGTACGATCTGTAAGCTTGAT 1461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15928 ATTTTCTTTTTCATTAATTTTAAATTTTATTTTATTAATTAATTTTATTT 15869
QY 1462 AAGTGTGACCTAGTGAGTTAAATTCGTGCTATTCAGTTGGAATCTTCAACC 1521
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15868 TATATTTTCTTTTATTAATTAATTTTATTTTATTTTATTTTATTAATTA 15809
QY 1522 TGCATTCGTAT-----GATCACTGGGGAATATTTGGGACATATTCGGAAC 1576
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15808 TTTTCTTTTATTTTATTAATTAATTTTATTTTATTTTATTTTATTTTAT 15749
QY 1577 TTCAATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1636
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15748 TGATATATTTTATTTATTAATTTTCTTTTCTTTTCTTTTCTTTTCTTTT 15689
QY 1637 TTTGTTTCTAGTCTATTTTGTAGAAATTTGACCTTTTGTGCTAAACTATTT 1696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15688 TTTTCTTTTAAATGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 15629
QY 1697 TTTTCTTGTACAGGCGCATGTGACCTTCATCAAGTACTGTTTCATGTGTAAC 1756
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15628 TTTTCTTGTACAGGCGCATGTGACCTTCATCAAGTACTGTTTCATGTGTA 15570
QY 1757 TAAATATGACTCTATTTGAGGTTATTTCTGCTGATCGATCTGTTTGTGAA 1816
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15569 TATCAAAATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 15510
QY 1817 TTGCTTGTGGGATTTATTTGAGAACTGCTGCTTACGACCACT 1863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 15509 TTTTCTTTTCTTTTATTAATTAATTTTCTTTTATTTCTATTT 15463

```

```

RESULT 15
US-09-443-041A-1
; Sequence 1, Application US/09443041A
; Patent No. 6465717

```

```

; GENERAL INFORMATION:
; APPLICANT: Famodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; PRIOR FILING DATE: 1999-11-18
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Zea mays
; US-09-443-041A-1

```

```

Query Match 0.98; Score 45.6; DB 4; Length 624;
Best Local Similarity 71.4%; Pred. No. 0.037;
Matches 60; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
QY 675 CGGAGACTGTACATTCCTCCGATCGTTACTTACGACATGATGTTATCGCTTC 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 140 CGGTACAGTGCACCTCGGGCGTGGTCACTACACCTTCATGCTCGGGCTCTCT 199
QY 735 GTCCACCTTTCGTCATTCCTCTGT 758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 200 GCCCGCCCTTCGTCATTCCTCTGT 223

```

```

Search completed: January 16, 2003, 21:26:43
Job time : 240 secs

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 13:15:32 ; Search time 12118 Seconds

(without alignments)
11717.494 Million cell updates/sec

Title: US-09-817-774-30

Perfect score: 4879

Sequence: 1 aactgtcctcctccacaa.....atctaaacaatttactg 4879

Scoring table: IDENTITY NUC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBml: *
1: gb_da: *
2: gb_hlg: *
3: gb_in: *
4: gb_om: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_da: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_hlg_hum: *
31: em_hlg_inv: *
32: em_hlg_other: *
33: em_hlg_mus: *
34: em_hlg_pin: *
35: em_hlg_rod: *
36: em_hlg_mam: *
37: em_hlg_vrt: *
38: em_sy: *
39: em_hlg_hum: *
40: em_hlg_mus: *
41: em_hlg_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4868	99.8	4880	8	AF239701
2	4750	97.4	105733	8	AF239701 Arabidops
3	251.8	5.2	1547	8	AC012561 Arabidops
4	237.4	4.9	1194	8	AY09589 Arabidops
5	224	4.6	1496	6	AY085063 Arabidops
6	224	4.6	1496	6	AY079111 Arabidops
7	224	4.6	1496	6	AR083377 Arabidops
8	224	4.6	1497	8	AT049398 Arabidops
9	135.6	2.8	125348	2	AP004850 Arabidops
10	135.6	2.8	178138	2	AP005514 Arabidops
11	75	1.5	143585	2	AC013349 Arabidops
12	69.2	1.4	9539	6	AX277889 Arabidops
13	69.2	1.4	9539	6	AX323566 Arabidops
14	68.8	1.4	6668	6	AX346599 Arabidops
15	67.2	1.4	166544	2	AC125605 Arabidops
16	66.8	1.4	70511	2	AC091083 Arabidops
17	66.6	1.4	7218	6	166494 Arabidops
18	66.4	1.4	156550	2	AC015830 Arabidops
19	65.6	1.3	1141	6	AX083744 Arabidops
20	65.6	1.3	157362	2	CNS08CAU Arabidops
21	65.4	1.3	86581	2	AC106497 Arabidops
22	65.4	1.3	176663	2	AC117848 Arabidops
23	64.6	1.3	114897	2	AP003624 Arabidops
24	64.2	1.3	34769	6	AX251533 Arabidops
25	63.8	1.3	6644	6	E23356 Arabidops
26	63.8	1.3	7372	6	E23357 Arabidops
27	63.8	1.3	7797	6	E23355 Arabidops
28	63.8	1.3	7996	6	E23359 Arabidops
29	63.8	1.3	14006	6	AX346860 Arabidops
30	63.8	1.3	349980	6	AX344559 Arabidops
31	63.8	1.3	349980	6	AX344570 Arabidops
32	63.8	1.3	349980	6	AX344571 Arabidops
33	63	1.3	210105	2	AC125948 Arabidops
34	62.8	1.3	83440	2	AC024285 Arabidops
35	62.8	1.3	170568	2	AC107605 Arabidops
36	62.4	1.3	184535	2	CNS05RCL Arabidops
37	62	1.3	538	8	AF417293 Arabidops
38	61.8	1.3	349980	6	AX344566 Arabidops
39	61.6	1.3	18598	6	AX345288 Arabidops
40	61.4	1.3	33068	3	AC115682 Arabidops
41	61.4	1.3	131271	2	AC015927 Arabidops
42	61.2	1.3	840	8	CNS0180K Arabidops
43	60.8	1.3	9864	6	AX345000 Arabidops
44	60.8	1.2	147962	2	AC130632 Arabidops
45	60.6	1.2	761	11	CNS06KMG Arabidops

ALIGNMENTS

RESULT 1
AF239701
LOCUS Arabidopsis thaliana sterol delta7 reductase (DWARF5) gene,
DEFINITION complete cds.
ACCESSION AF239701
VERSION AF239701.1 GI:7542560
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidops;
1 (bases 1 to 4880)
REFERENCE
Choe,S., Tanaka,A., Noguchi,T., Fujioke,S., Takatsuto,S.,

Pred. No. is the number of results predicted by chance to have a

TITLE Ross, A.S., Tax, F.E., Yoshida, S. and Feldmann, K.A.
Lesions in the sterol delta reductase gene of Arabidopsis cause
dwarfism due to a block in brassinosteroid biosynthesis
JOURNAL Plant J. 21 (5), 431-443 (2000)
MEDLINE 20223137
PubMed 10758495
REFERENCE 2 (bases 1 to 4880)
AUTHORS Choe, S.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2000) Plant Sciences, University of Arizona,
Forbes hall 303, Tucson, AZ 85721, USA

FEATURES

source
1. 4880
/organism="Arabidopsis thaliana"
/cultivar="Ws-2"
/db_xref="taxon:3702"
/chromosome="1"
/map="linked to nga280"
/gene="DMARF5"
/note="linked to nga280"
/gene="DMARF5"
join(<671..756,1049..1271,1340..1395,1504..1606,
1711..1776,2174..2287,2418..2544,2774..2874,3259..3384,
3494..3558,3635..3716,3806..3888,3981..4047)
/gene="DMARF5"
/product="sterol delta7 reductase"
join(671..756,1049..1271,1340..1395,1504..1606,1711..1776,
2174..2287,2418..2544,2774..2874,3259..3384,3494..3558,
3635..3716,3806..3888,3981..4047)
/gene="DMARF5"
/note="sterol biosynthetic protein"
/codon_start=1
/product="sterol delta7 reductase"
/protein_18="AAP63498.1"
/db_xref="GI:7542361"
/translation="MAETVHSPVITYVSMLSLAFCEPPYVILMTYVHDDGVTQTF
GFENGVQGLINIPRETLIAKIIICYGAFEAIILOLLPGKVEEIPISAGNRPVY
KANGIAAFVLAIVLGLMVEINPAIVDHLGEISALIFGSEICVLIKHVA
PSSDSGSCGNLIDFYMGMELPRIGKSPDIKFTKCRFEMSMALVAATYCKOYE
INGKVSMDMLVTILMAYVTKFPMMEGWNVDIAHRAQFVIGCGCLVWVSVT
SPKATLVNHPVELSGTQLAIVLIVAGIICITVYNPCDROFRRRTNCKLVMSAPSK
IVASTTTSGETKSLSLDSGKMWGLAFHFIHVPPELSAFETVYVATLDNPLAYVYVF
LTLLEFDRAKDDRCRSKRYKMWKLYCEKVKYKRIIPETV"

BASE COUNT 1409 a 889 c 922 g 1660 t
ORIGIN

Query Match 99.8%; Score 4868; DB 8; Length 4880;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4879; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 1 AACTGTCTCTCCACACACAGATCTCTCTCCGCTCTACGATTTGTTAAG 60
|||||
Db 1 AACTTGTCTCTCCACACACAGATCTCTCTCCGCTCTACGATTTGTTAAG 60
OY 61 ATGTGATTCGTCGAATAAATCTGTTCTTCTATCATCATCATCAGCAGCTA 120
|||||
Db 61 ATGTGATTCGTCGAATAAATCTGTTCTTCTATCATCATCATCAGCAGCTA 120
OY 121 ATTGCAATCTTAAACATCTAAATTTCTGGAATGAAGAAAAAAGTGAAGCAAAA 180
|||||
Db 121 ATTGCAATCTTAAACATCTAAATTTCTGGAATGAAGAAAAAAGTGAAGCAAAA 180
OY 181 GAATTCACACAGCATCAGAAAACTAGATCTAGTTTCCGCGGAGTTACCAAAAAACAT 240
|||||
Db 181 GAATTCACACAGCATCAGAAAACTAGATCTAGTTTCCGCGGAGTTACCAAAAAACAT 240
OY 241 AGAAAAATCGATCTTACACGATTTCAACCAATTTCTGAGAAATTTTCTTCCAGAGCAAG 300
|||||
Db 241 AGAAAAATCGATCTTACACGATTTCAACCAATTTCTGAGAAATTTTCTTCCAGAGCAAG 300
OY 301 CCAATTCCTGCGAGTAAGAAAGAAATGTAACATATGATCGCAATCAAGTGAAGAGA 360
|||||
Db 301 CCAATTCCTGCGAGTAAGAAAGAAATGTAACATATGATCGCAATCAAGTGAAGAGA 360

OY 361 TTGATTTTGTAGTACGATTCATCATCTATGTTGAATACAGAGATGACGAAGATCT 420
|||||
Db 361 TTGATTTTGTAGTACGATTCATCATCTATGTTGAATACAGAGATGACGAAGATCT 420
OY 421 TTGAGGTTGAGAGAGAGAGAGAGTAAAGATTGGGATCAATCAAAAACGATGGGAGA 480
|||||
Db 421 TTGAGGTTGAGAGAGAGAGAGAGTAAAGATTGGGATCAATCAAAAACGATGGGAGA 480
OY 481 AAGTCTCAAGGTTAAATTTGGAATTAAGAAAAAAGTGGAGAGCAGTGCAGCTGG 540
|||||
Db 481 AAGTCTCAAGGTTAAATTTGGAATTAAGAAAAAAGTGGAGAGCAGTGCAGCTGG 540
OY 541 ATTATCTGAGTTTCAAGTACAGTACCTGTAGTTTCAATCTTCAATCTTCAATCTG 600
|||||
Db 541 ATTATCTGAGTTTCAAGTACAGTACCTGTAGTTTCAATCTTCAATCTTCAATCTG 600
OY 601 ACTGTGTTGATTTGTCACACACATCAGAAATCTGAGGCTTTGGCCGAGAGAGAGA 660
|||||
Db 601 ACTGTGTTGATTTGTCACACACATCAGAAATCTGAGGCTTTGGCCGAGAGAGAGA 660
OY 661 GAAGAAAGAAATGGCGAGAGCTGATATCTCCGATCTCCGATCTTACGATCATGTTATC 720
|||||
Db 661 GAAGAAAGAAATGGCGAGAGCTGATATCTCCGATCTCCGATCTTACGATCATGTTATC 720
OY 721 GCTTCTGCGCTTCTGTCACCTTCTGTCATCTCTGTAAGCTCATCAATTTCTGATTCG 780
|||||
Db 721 GCTTCTGCGCTTCTGTCACCTTCTGTCATCTCTGTAAGCTCATCAATTTCTGATTCG 780
OY 781 CAACTCTTATTTCTAGTTCACATCAGATCGCACACTTTCTGCTTCAATTTCA 840
|||||
Db 781 CAACTCTTATTTCTAGTTCACATCAGATCGCACACTTTCTGCTTCAATTTCA 840
OY 841 GCGAGAGTGCATAGCCGATATCTTAAATCTTCCCATTAAGTGAAGCTGATTAAG 900
|||||
Db 841 GCGAGAGTGCATAGCCGATATCTTAAATCTTCCCATTAAGTGAAGCTGATTAAG 900
OY 901 TCTTTAATGAAGAAAGGATTCATGATTTCTTCTGCGCAATGTGATTTCAATTT 960
|||||
Db 901 TCTTTAATGAAGAAAGGATTCATGATTTCTTCTGCGCAATGTGATTTCAATTT 960
OY 961 TTAGAAACAAAGGTTAGATTTAAAGTACAAATATCATATTAGTGAATTAAGTAA 1020
|||||
Db 961 TTAGAAACAAAGGTTAGATTTAAAGTACAAATATCATATTAGTGAATTAAGTAA 1020
OY 1021 TCACAGTTTCATTTCTTATGATGACAGATGTACACAAATGTTTCAGAGATGTTCTG 1080
|||||
Db 1021 TCACAGTTTCATTTCTTATGATGACAGATGTACACAAATGTTTCAGAGATGTTCTG 1080
OY 1081 TTACTCAGACCTTTGGCTTCTTTGGGAGAAATGAGATTCAGAAAGTATCAATATGGC 1140
|||||
Db 1081 TTACTCAGACCTTTGGCTTCTTTGGGAGAAATGAGATTCAGAAAGTATCAATATGGC 1140
OY 1141 CAAGACCACCTTATGATTTGTTGAATTAATTTCTATGAGAGCTATTTG 1200
|||||
Db 1141 CAAGACCACCTTATGATTTGTTGAATTAATTTCTATGAGAGCTATTTG 1200
OY 1201 TTCACTGCTCTCTGCTGTTAAAGAGTGAAGGTTCCAAATATCTCCAGCGGAAACGAC 1260
|||||
Db 1201 TTCACTGCTCTCTGCTGTTAAAGAGTGAAGGTTCCAAATATCTCCAGCGGAAACGAC 1260
OY 1261 CAGTTTACAGATTTGTTCAATTTAGTCTGTGTGTGGAAGTGAAGTGAAGTGAAGTGA 1320
|||||
Db 1261 CAGTTTACAGATTTGTTCAATTTAGTCTGTGTGTGGAAGTGAAGTGAAGTGAAGTGA 1320
OY 1321 TACGTTCTTTCTTGGCAGGCAATGCTGCTTACTTGTGACACTAGCAACCTA 1380
|||||
Db 1321 TACGTTCTTTCTTGGCAGGCAATGCTGCTTACTTGTGACACTAGCAACCTA 1380
OY 1381 TCTTGGTCTTGTGTTGTAAGAAATGTTCCGATGATCTTGGTTTGAATTAATTT 1440
|||||
Db 1381 TCTTGGTCTTGTGTTGTAAGAAATGTTCCGATGATCTTGGTTTGAATTAATTT 1440
OY 1441 TACGATCTTAAAGCTTGAATTAAGTGTTCACACTATGAGTTAAATTTCTGTGGTTAT 1500
|||||

|||||
Db 1441 TACGATCTTAAAGCTTGAATTAAGTGTGACCTATGTGAGTTAATTTCTGTGCTATT 1500
QY 1501 CAGGTTTGAATCTTCAACCCCTGCAATTTGCTATGATCAGCTGGCTGAATATTTTGGC 1560
Db 1501 CAGGTTTGAATCTTCAACCCCTGCAATTTGCTATGATCAGCTGGCTGAATATTTTGGC 1560
QY 1561 ACTAATATTCGGAAGCTTCATATTTTGTGTTTGTGTACATAAAGTAAGTGTATG 1620
Db 1561 ACTAATATTCGGAAGCTTCATATTTTGTGTTTGTGTACATAAAGTAAGTGTATG 1620
QY 1621 CATGAATGCTGTCTTTGTTTCTTATGCTCTATTTTGTAGAAATTTGACCTTTTGTG 1680
Db 1621 CATGAATGCTGTCTTTGTTTCTTATGCTCTATTTTGTAGAAATTTGACCTTTTGTG 1680
QY 1681 TTGCTAAACCTATTTTCTTCTGACAGGGCCATGTGACCTCTATCAAGTACTCT 1740
Db 1681 TTGCTAAACCTATTTTCTTCTGACAGGGCCATGTGACCTCTATCAAGTACTCT 1740
QY 1741 GGTTCATGTGTACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
Db 1741 GGTTCATGTGTACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1800
QY 1801 CTCATTTCTTTTGAATTTGCTTTGTTGGATTTATTTGAGAACCTGTGCTTACTAGCA 1860
Db 1801 CTCATTTCTTTTGAATTTGCTTTGTTGGATTTATTTGAGAACCTGTGCTTACTAGCA 1860
QY 1861 TCTCCATATAATTTTGGCTTTTCTGACATTTGAGGTCACATTTTCACTAAAGT 1920
Db 1861 TCTCCATATAATTTTGGCTTTTCTGACATTTGAGGTCACATTTTCACTAAAGT 1920
QY 1921 TCCATATTTAGTAGACCTGTAAACGGAATTTCTTAGCAGTATTTATCTTTTATGAT 1980
Db 1921 TCCATATTTAGTAGACCTGTAAACGGAATTTCTTAGCAGTATTTATCTTTTATGAT 1980
QY 1981 TTAGTACAGAAACCTAGTTTGTGTGATGTCTGTGCTGAAGAAAGAAAGAAAGT 2040
Db 1981 TTAGTACAGAAACCTAGTTTGTGTGATGTCTGTGCTGAAGAAAGAAAGAAAGT 2040
QY 2041 TGCAGTATTTATTAACATATCAATGATGATCTCTGACACATTTCTCTTTTGTCT 2100
Db 2041 TGCAGTATTTATTAACATATCAATGATGATCTCTGACACATTTCTCTTTTGTCT 2100
QY 2101 CCAACTATATCAATTAACCAATTCGAATTTGTGGAATCTTTCTACAGTGTTTTTCGCG 2160
Db 2101 CCAACTATATCAATTAACCAATTCGAATTTGTGGAATCTTTCTACAGTGTTTTTCGCG 2160
QY 2161 TCGTGTCTTTCAGGGGCTGAGAGTGTGACCTCGAATTTGAGAGCTTTGACATCAAGT 2220
Db 2161 TCGTGTCTTTCAGGGGCTGAGAGTGTGACCTCGAATTTGAGAGCTTTGACATCAAGT 2220
QY 2221 GTTACTAATTTGAGATTTGGAATGATGCTTGGGAGCTTCTGACGATCGATAGCAT 2280
Db 2221 GTTACTAATTTGAGATTTGGAATGATGCTTGGGAGCTTCTGACGATCGATAGCAT 2280
QY 2281 AAAACAGATTTTCTGATATTTCTATGTAATTAATTAATTAATTAATTAATTAAT 2340
Db 2281 AAAACAGATTTTCTGATATTTCTATGTAATTAATTAATTAATTAATTAATTAAT 2340
QY 2341 TGGTCTTCTTAAGTCTTATTTCTATGAGAGCTTCACTCTTCTGATTTGTGATTCCT 2400
Db 2341 TGGTCTTCTTAAGTCTTATTTCTATGAGAGCTTCACTCTTCTGATTTGTGATTCCT 2400
QY 2401 GTACTTCTTAAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2460
Db 2401 GTACTTCTTAAACAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2460
QY 2461 CCAATCTGATGCTGTGTATGACAAAATTTCTTGTGTGGAAGCTGTTTATGGAACA 2520
Db 2461 CCAATCTGATGCTGTGTATGACAAAATTTCTTGTGTGGAAGCTGTTTATGGAACA 2520
QY 2521 CCATGACATTTGACATGACCGAGGATGCTTGTGATCATATTAATTAATTAATTAAT 2580
Db 2521 CCATGACATTTGACATGACCGAGGATGCTTGTGATCATATTAATTAATTAATTAAT 2580

Db 2521 CCATGACATTTGACATGACCGAGGATGCTTGTGATCATATTAATTAATTAATTAAT 2580
QY 2581 ACATCTCTATATAGGTTACAGAGAAAGAAAGCCAGTATTAATTAATTAATTAATTAAT 2640
Db 2581 ACATCTCTATATAGGTTACAGAGAAAGAAAGCCAGTATTAATTAATTAATTAATTAAT 2640
QY 2641 GTTGGCTATTAATTTCTTCCAGAGTCTGTTTATATAGGTTAAGGAGCGCATTTATTC 2700
Db 2641 GTTGGCTATTAATTTCTTCCAGAGTCTGTTTATATAGGTTAAGGAGCGCATTTATTC 2700
QY 2701 AGATTAATTTCTTACGCTTTAATCAATGAATTAATTAATTAATTAATTAATTAATTAAT 2760
Db 2701 AGATTAATTTCTTACGCTTTAATCAATGAATTAATTAATTAATTAATTAATTAATTAAT 2760
QY 2761 TATGATTTCTTACGCTGATTTATATATCTGAGGTTGTCTAGTGTGCTGCTTCTGTC 2820
Db 2761 TATGATTTCTTACGCTGATTTATATATCTGAGGTTGTCTAGTGTGCTGCTTCTGTC 2820
QY 2821 TACACTTCTCAGGAGTGTACCTGTGAGACCAACCCGCTGAGCTGGAATCAGTGCAT 2880
Db 2821 TACACTTCTCAGGAGTGTACCTGTGAGACCAACCCGCTGAGCTGGAATCAGTGCAT 2880
QY 2881 AATTTGAATTTGTCTAGAAAGAAAGAAATCGTCTTGTGTGTTAATCCATGAGAACTT 2940
Db 2881 AATTTGAATTTGTCTAGAAAGAAAGAAATCGTCTTGTGTGTTAATCCATGAGAACTT 2940
QY 2941 TTAGATTAACCTTTTATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT 3000
Db 2941 TTAGATTAACCTTTTATTAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGT 3000
QY 3001 CCAATTTCTAGAAATTTGCACTTAACCTTGTATTTGATTAATTTGATTTCTTCAATTTCA 3060
Db 3001 CCAATTTCTAGAAATTTGCACTTAACCTTGTATTTGATTAATTTGATTTCTTCAATTTCA 3060
QY 3061 AACAGACGAGAGAGAGTAAGTGTAGATTTGTATCAATTTGCTGTGATTTATGCTC 3120
Db 3061 AACAGACGAGAGAGAGTAAGTGTAGATTTGTATCAATTTGCTGTGATTTATGCTC 3120
QY 3121 ATGATTTGATTTGCTCAGATATCAGTTTCTAGTATGATAGAAAGCTGTTATATTAAC 3180
Db 3121 ATGATTTGATTTGCTCAGATATCAGTTTCTAGTATGATAGAAAGCTGTTATATTAAC 3180
QY 3181 ACCAGTATGATACGATCAATTCGAGCAATTAATGCTATCTGCTTACCTTAAGCTTACAT 3240
Db 3181 ACCAGTATGATACGATCAATTCGAGCAATTAATGCTATCTGCTTACCTTAAGCTTACAT 3240
QY 3241 ATCCGCTTTTGTGCTGAGTGTGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300
Db 3241 ATCCGCTTTTGTGCTGAGTGTGCAATTAATTAATTAATTAATTAATTAATTAATTAAT 3300
QY 3301 ATTAACATATGATGATGATTAAGAAAGCAAGGTTTCAAGAGGACAAACGGGAATGTTG 3360
Db 3301 ATTAACATATGATGATGATTAAGAAAGCAAGGTTTCAAGAGGACAAACGGGAATGTTG 3360
QY 3361 GTTGGGAGAGAGCCCGTCAAAAGGTTGATCATTAATTAAGCCCACTAGTGTGAGAC 3420
Db 3361 GTTGGGAGAGAGCCCGTCAAAAGGTTGATCATTAATTAAGCCCACTAGTGTGAGAC 3420
QY 3421 TTAGCTATATTTCTTTTCAAGACCTTAACATTTTGTATGATTAATTAATTAATTAAT 3480
Db 3421 TTAGCTATATTTCTTTTCAAGACCTTAACATTTTGTATGATTAATTAATTAATTAAT 3480
QY 3481 CGGTTTGTGAGATTTGCGCTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3540
Db 3481 CGGTTTGTGAGATTTGCGCTGCTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3540
QY 3541 TCTCTTAACGCTGTGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3600
Db 3541 TCTCTTAACGCTGTGATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3600
QY 3601 ACCTTGTTCATCATATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
Db 3601 ACCTTGTTCATCATATGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660

OY	3661	ATGTCCTGGAGACTCTTAAGAGCTTCTCTTGACCGCTACCGGCTCTCTCATTAACGTA	3720
Db	3661	ATGTCCTGGAGACTCTTAAGAGCTTCTCTTGACCGCTACCGGCTCTCTCATTAACGTA	3720
OY	3721	AATACATACATCTATTGGCTATATTGCAAGCTCTTCTGCAAGTAAGACGACTTAACA	3780
Db	3721	AATACATACATCTATTGGCTATATTGCAAGCTCTTCTGCAAGTAAGACGACTTAACA	3780
OY	3781	TATAAATCATCTTGGTTAAACAGATCTTGCGATCTCTACGTCATATTCTCCACCCT	3840
Db	3781	TATAAATCATCTTGGTTAAACAGATCTTGCGATCTCTACGTCATATTCTCCACCCT	3840
OY	3841	TCTTCTCTTTGATGAGCCCAAGAGACGATGACCGATGCGATCAAGTAAGACCCAA	3900
Db	3841	TCTTCTCTTTGATGAGCCCAAGAGACGATGACCGATGCGATCAAGTAAGACCCAA	3900
OY	3901	TCTAAATGATTTATTATTACACATCTCTTGATCTTACTCTATTTGCTGCAATATTGCG	3960
Db	3901	TCTAAATGATTTATTATTACACATCTCTTGATCTTACTCTATTTGCTGCAATATTGCG	3960
OY	3961	GGTTTTGTTTTGGTAAAGGATAGGAAATATTGGAAGCTGATTTGGAG- AAGTCAAA	4019
Db	3961	GGTTTTGTTTTGGTAAAGGATAGGAAATATTGGAAGCTGATTTGGAGTATGCAAAAGTCAAA	4020
OY	4020	TACAGATCATTCGGGAATTTATGTATTAAACAATCTGTGTTCTCATTTTCTACT	4079
Db	4021	TACAGGATCATTCGGGAATTTATGTATTAAAGAGCTGTGTCTCATTTTCTACT	4080
OY	4080	TATTACGTTAATTGCAAGCTTTGGATCATTAAGAAGCCGGCCAAACCAAAATGCAAT	4139
Db	4081	TATTACGTTAATTGCAAGCTTTGGATCATTAAGAAGCCGGCCAAACCAAAATGCAAT	4140
OY	4140	TGATGCGATAGACATCTCTTGGCTGATTTGTATGCTATAGTTTTCAAACTCTAGTAC	4199
Db	4141	TGATGCGATAGACATCTCTTGGCTGATTTGTATGCTATAGTTTTCAAACTCTAGTAC	4200
OY	4200	GCTATGATTCCTCGATTAATCAAAAGTAGCCGCGTTTCTAATTTTATTTGA	4259
Db	4201	GCTATGATTCCTCGATTAATCAAAAGTAGGCCCTGCGTTTCTAATTTTATTTGA	4260
OY	4260	TATTTTGATTTACTGCTTCTTAATGATCAATAATAGTTGATTCATCATCTAATATTTT	4319
Db	4261	TATTTTGATTTACTGCTTCTTAATGATCAATAATAGTTGATTCATCATCTAATATTTT	4320
OY	4320	ACTATCAATGAAGCTTTTAAATGAAGGACCTAATGAACCTGAAGCTGTAA	4379
Db	4321	ACTATCAATGAAGCTTTTAAATGAAGGACCTAATGAACCTGTAA	4380
OY	4380	TACATGAATTAATCAAAAGATTGATAGTCTTCTTACGTTTTCATCACAAAGACTC	4439
Db	4381	TACATGAATTAATCAAAAGATTGATAGTCTTCTTACGTTTTCATCACAAAGACTC	4440
OY	4440	ACCGACATTATCAAGTGTACTTCACTTTAGACTTAATGTGTCTGAGAGATCCACA	4499
Db	4441	ACCGACATTATCAAGTGTACTTCACTTTAGACTTAATGTGTCTGAGAGATCCACA	4500
OY	4500	AACACTATTACAGAGTTTTCACAAATTTGATTCCTAATTTTAAATTTCTTTGCTCA	4559
Db	4501	AACACTATTACAGAGTTTTCACAAATTTGATTCCTAATTTTAAATTTCTTTGCTCA	4560
OY	4560	AACATACTTTCATTCAACTTTAAACGAGATTTGCCACAGTGAGAGTATACAAAA	4619
Db	4561	AACATACTTTCATTCAACTTTAAACGAGATTTGCCACAGTGAGAGTATACAAAA	4620
OY	4620	AGCTACGAGACAATGCTCAAAACAAACATTAAGAAAGGAGAGCTCGAAGCTCAAC	4679
Db	4621	AGCTACGAGACAATGCTCAAAACAAACATTAAGAAAGGAGAGCTCGAAGCTCAAC	4680
OY	4680	CGAGTAGCAGATCCAAAGTTGCTAAACAGAACCATCTATTCTCTGAGCGGTCTTC	4739
Db	4681	CGAGTAGCAGATCCAAAGTTGCTAAACAGAACCATCTATTCTCTGAGCGGTCTTC	4740

OY	4740	GTTCACAGCCTTCGACGAAATTTGATTCATGTTGGATTCATGACATCTACTACTTC	4799	
Db	4741	GTTCACAGCCTTCGACGAAATTTGATTCATGTTGGATTCATGACATCTACTACTTC	4800	
OY	4800	AAAGTTTAAATACAAAACCTTACCATTCTCTATTTGTTAGTCAAGCTAACAAATTTTC	4859	
Db	4801	AAAGTTTAAATACAAAACCTTACCATTCTCTATTTGTTAGTCAAGCTAACAAATTTTC	4860	
OY	4860	ATCTAAAAACAATTTACTG	4879	
Db	4861	ATCTAAAAACAATTTACTG	4880	
RESULT 2	AC012561	105733 bp	DNA linear	PLN 29-JUL-2000
LOCUS	AC012561	Arabidopsis thaliana chromosome I BAC F11P12 genomic sequence,		
DEFINITION	complete sequence.			
ACCESSION	AC012561			
VERSION	AC012561.2	GI:6227026		
KEYWORDS	HTG.			
SOURCE	Arabidopsis thaliana.			
ORGANISM	Arabidopsis thaliana.			
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
REFERENCE	1 (bases 1 to 105733)			
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altai,H., Arujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,A., Liu,S., Luros,S., Schwartz,J., Shinn,P., Tortum,M., Vyotskaia,V.S., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.			
JOURNAL	Unpublished			
REFERENCE	2 (bases 1 to 105733)			
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altai,H., Nguyen,M., Lam,B., Southwick,A., Bel,O., Buehler,E., Chin,C., Chou,J., Choi,E., Dunn,P., Gonzalez,A., Howing,B., Kim,C., Koo,T., Lee,J.M., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharly,N., Pham,P., Sakano,H., Schwartz,J., Shinn,P., Thavari,A., Tortum,M., Vaysberg,M., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.			
JOURNAL	Direct Submission			
TITLE	Submitted (30-OCT-1999) DNA Sequencing and Technology Center.			
JOURNAL	Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA			
REFERENCE	3 (bases 1 to 105733)			
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altai,H., Arujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S., Luros,S., Schwartz,J., Shinn,P., Tortum,M., Vyotskaia,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.			
JOURNAL	Direct Submission			
TITLE	Submitted (04-NOV-1999) DNA Sequencing and Technology Center.			
JOURNAL	Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA			
REFERENCE	4 (bases 1 to 105733)			
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altai,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A. and Davis,R.W.			
JOURNAL	Direct Submission			
TITLE	Submitted (26-JUL-2000) DNA Sequencing and Technology Center.			
JOURNAL	Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA			
REFERENCE	5 (bases 1 to 105733)			
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F., Altai,H., Nguyen,M., Lam,B., Southwick,A., Ecker,J., Theologis,A. and Davis,R.W.			
JOURNAL	Direct Submission			
TITLE	Submitted (29-JUL-2000) DNA Sequencing and Technology Center.			
JOURNAL	Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA			
COMMENT	On Nov 4, 1999 this sequence version replaced gi:6143852. Bases 100173-105733 of TGF BAC clone F11P12 overlap with bases			

FEATURES
source
1-5561 of IGF BAC clone F1413, gb|AC007980.
Location/Qualifiers
1. .105733
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F11F12"
479. .1945
/gene="F11F12.1"
join(479. .543,626. .671,917. .1025,1199. .1290,1370. .1432,
1527. .1645,1733. .1808,1889. .1945)
/gene="F11F12.1"
/note="Unknown protein; similar to CG4603 gene product
from Drosophila melanogaster"
/codon_start=1
/protein_id="AAF87868.1"
/db_xref="GI:9454545"
/translation="MEGITVRYVPSDSCLENAIGYVMDKKNKAPBLRQVIAAVA
SNKRYNEAFVGLNEGCAMILNPKMGALIELSLADYGRLEIAADIQSRCDLY
GQRNYDERVMLYDGLHYDALALSPFEGAEEDFDMTIVPGKDRSIGSIEGLALNV
KQOQRSTYDTPANFTLRGVCQIGVIGQKEAVEHAQATGHVNFQEK"
complement(2646. .5645)
/gene="F11F12.2"
complement(join(2646. .4253,4769. .4897,5205. .5645))
/gene="F11F12.2"
/note="Unknown protein"
/codon_start=1
/protein_id="AAF87869.1"
/db_xref="GI:9454546"
/translation="KAVTGKPYSTATLPAPESDKNLSHPDLRALIQRATVTKRRAR
NPSLTRQSRSGGRRSRPEPLKMKVEDNKRKSGYEDDDYEDDHNQVARSFTT
RRDRRIARPSGVRKRLAQLWRLQVDPASSGGERKGEGLGPOGNGYMPVLYH
HSKPSGGSQNRKRONSTIATKNGFLCKLPSEMPFPMSAMEGATKMDPVOLDME
VHDTYSNMRIDQOVAVSLVSLAEALREAHRIEDLESERHKKEQTLRYVSE
ERAWSRREHEVRAIIDDKMDNREKTRORLEIVNKLVELADSLAKRYAMOD
YERKARRELEVCDELAKETIGEDALIEALKRESMSREEVDERRELQOAEVRE
ERYOMKLIQAKVLEERYSQMKLVGLDLSFRLSRIVTVKREARELRLQOAEVRE
IOIKETFPYANPDIVAVFEEMNLGEHDEMEXVAYSPISHSKHYVSLAAM
MNGKGRHSAYTHQNGDIEDDSGMEVSHLEEOGSSYSPGDSIPSYNNKNNHNS
ASSGTESELGKWDMDTMTPTTEISEYCSIPRRSKKVSITAKMTSGASNDROSN
KYSMEGNGKGRVSDNRKSSAGMVSFDRSSSGGSPMDLVGOMNSSPESANHPHVN
RSGMKCTEMPRAQKSSLSKLIEARIESQVOKHVLKORI"
12737. .13258
/gene="F11F12.3"
12737. .13258
/gene="F11F12.3"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAF87870.1"
/db_xref="GI:9454547"
/translation="MSIKLTLCALIEFLINSLIHLVHGHNDLNTTSMKSHIKAT
TINMGKPPMKPCPKPICRSGPSEARMRCRCNOGVYLSDPHRCRPFCECFALSSC
DGGCVDTNTPNSCGCGNCSGACERCMGYAPSSQPKRRIRKHFHRPPPS
PSKLTNYDDHDE"
18752. .19429
/gene="F11F12.4"
/note="ALERT-3"
18752. .19429
/gene="F11F12.4"
/note="cDNA clone gb|AB008105"
/codon_start=1
/product="ethylene responsive element binding factor 3"
/protein_id="AAF87871.1"
/db_xref="GI:9454548"
/translation="MRGRSSAVAGPTVVAALNGSVKEIRRGYRKKRPGFAAER
DPMKRVAVLGFDSAEARAYDSARARLRPKAKTNFPIDSSPPPNLFTNDIRN
QNOVYDPPMDHRLFTDHOQEPYINRPTSSMSSTVESFSQPRPTTKMPATTKRPR
TPPVPEDCSDSSSVIIDDDDIASSRRRNPDPDLNFPPLDCVDFNGADLL
HCTDLRL"
23462. .25377
/gene="F11F12.5"
join(23462. .22896,24283. .24732,24817. .25156,25241. .25377)
/gene="F11F12.5"

/note="Unknown protein"
/codon_start=1
/protein_id="AAF87872.1"
/db_xref="GI:9454549"
/translation="MADIDVAGKEPLINHENNNNTKIFNRCSVHODDELHSFRKYL
RMVCVSHSPWTAISMTMFIIVFTLVPAISHPLACADCSYSRPYDSVOLSTLS
VATVFLCTREVSRYGLRFLPEFKLDESEETVRNTNOLNLSHVSFVPCPS
MASAKIMWVASGSRIPGLGNVSDVACIMELCSMLTPRTYVLEVCYFLRCLH
OILRQDAKFLFOIDSDVGSILSEHLIRRLRITISHRSFTICLLIVYGSFSL
LITKATYEVNIRAGELALCSMTLVALLILLSASAKITKRAQAVICLAKMVCAT
LESFOQVESPDQVETPTLVARNNNNNVHDVYVLTESDSEDEGDENDDNDII
PVAFSTMSFQRQALVSYFENNAGITVYGTFLDRGLHTIFELSLVLMIGKTI
GIS"
complement(26026. .28494)
/gene="F11F12.6"
complement(join(26026. .27287,27696. .27872,27948. .28052,
28176. .28494))
/gene="F11F12.6"
/note="Hypothetical protein"
/codon_start=1
/protein_id="AAF87873.1"
/db_xref="GI:9454550"
/translation="MDVAISDVVYNLESVEEKDEETVEESPVKKKRPDEENVRY
AEIVLVLSALKIRKRGKPELEIDLMEAKSLVDMCOEFTPDIDIGVADIGVIED
LGFNDRSEISLYVIVLDGIEKSSSIYIYISRDVATYNNVSMANHPNSEMKAS
ASSVNASGSHVREASGITQARMRPOKSLDHTGSGPAPVPAKNGFNGTTMSAP
HSSVTISFGTASVSKVYVPSRSISDSFPPRPSQTPPGAFPMKATYGGTSSPPG
NNNHAEIKLILIKVLOPRAKONLNNPSPREYMSKAMTCQICQGTINEIFVLIDC
EKGYHLKLAHANI KGVPSSEMHCSRQVLNKGSPFKYGRVRSATTKMSSSTAE
VQPAEKGVGMQDKVSEGMPHLETAPTDGSAEQTEVEDVAMNPIVEKAMSEV
EAEGAIPIYIEAEDGAMNPIVEKAMSOIYAEADAINOAVDANFOAPNOQADAS
DDSPVAPHSHTLNPPELEKEVWRKATPEVSVAQDADCKSKITASSIOEBLSAQ
TENSQTQPSQSDTDHSQOQTTHVNEBEKASQAEENSPTQPSQNTDHPYA"
complement(32526. .34856)
/gene="F11F12.7"
complement(join(32526. .33202,33287. .33961,34148. .34856))
/gene="F11F12.7"
/note="Putative protein kinase; similar to receptor
protein kinases"
/codon_start=1
/protein_id="AAF87874.1"
/db_xref="GI:9454551"
/translation="MRNDEPPTLACNTALKKNLPSCIFITIFISVLCPYMASGVVP
DSADCLIREKDTLNGSESPMDLSPGCGNANPFGVCSYVWGLDEMGILG
KLINDPLVPMKNLTISTPMNNPNPMPQYKRTSLKSLYSNNRSEGLIPADFLGM
PLKLILANNAFRGTTIPSSLASPLMLELRLNNOFGQIPSFQOKDLKLSAEPEND
LDGPTPELRMDPGSFANGKGLDAPLSGSSSPGVVAVPVPSPKSTSPPTGK
AGSFYTLAILIVIGIILVIALVFCVQSRRRRLSLAYSSAGKERIESYTHQSTN
KNKPDAESVNTTRGSMDDPGRLFLFYDIDIOFLDLBRASAEVIGSTFGASYKA
AISGQTLVVRKYKMMNVGDEFEHMRRLRGRNLNHNIIPIVAYYRRREKILVTEF
MNSSIASHLAHNNSAGLDWTRIKIKYAKGSLYLPDELPTITPHGMKSNLYL
DDSEPLITDIALRPMASSEIAHNPATYKSPERPSKQITTKTDVWFGVLLV
LTGRPENTLYOGYDNNSLVLTWNVNVEKKTVDYDKERKGNKAKAMINLKG
LRCEEERERMDREVEVEMVLEBESDDFGSMHRTGTHNNVSSMLDDDFG
SMNR"
38049. .39629
/gene="F11F12.8"
38049. .39629
/gene="F11F12.8"
/note="scarecrow-like 5 protein; scarecrow-like 5 protein
cDNA clone, gb|AF036302. putative transcription factor"
/codon_start=1
/product="putative transcription factor"
/protein_id="AAF87875.1"
/db_xref="GI:9454552"
/translation="MEATQKHMIDEGSMFYHOPSSYKQMDLSVQTFDYSCTLESSSG
TKSHPCLNKNNKSSSTSFSSNEPISQANNNNLSRNNHSPENNNSPLSGSSATNT
NETELSLMKLRLPAAQMEPDVNDYNNQGGCOOHGVSSAMYSRMEMIIRGLKGVL
YCAKAYENDLEMTDMLISQLOQMSVSEBPYGRLAGVMEDEIVARLASGSSYTKA

Query Match 97.4%; Score 4750; DB 8; Length 105733;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 4853; Conservative 0; Mismatches 25; Indels 8; Gaps 8;

QY	1	AAC	TTGCTCC	TCCTCC	CACAAACAGCAATCTC	CTTC	CCGGCTGCTACGAAATTC	GGTAAAG	60
Db	90803	AAC	CTGCTCC	TCCTCC	CACAAACAGCAATCTC	CTTC	CCGGCTGCTACGAAATTC	GGTAAAG	90861
QY	61	ATG	TTGATTCG	CGAGAT	AAATTAATCT	CTTCTCT	CTATCATCATCATCATCAGGACCTA	120	
Db	90862	ATG	TTGATTCG	CGAGAT	AAATTAATCT	CTTCTCT	CTATCATCATCATCATCAGGACCTA	90921	
QY	121	ATT	GCAATCTCT	TTAAACAT	CTAATAAT	TTTCTG	GAATGAAAAAACTGAAACCCAAA	180	
Db	90922	ATT	GCAATCTCT	TTAAACAT	CTAATAAT	TTTCTG	GAATGAAAAAACTGAAACCCAAA	90981	
QY	181	GAA	TTTCAACACAGCAT	CACAAAAC	TAGATCT	GTGGTT	CGCGGAGTTACCAAAAACAT	240	
Db	90982	GAA	TTTCAACACAGCAT	CACAAAAC	TAGATCT	GTGGTT	CGCGGAGTTACCAAAAACAT	91041	
QY	241	AGAAAAAAT	CGACTT	CTACGCAAT	TCAACCAAT	TTCTGAGA	ATTTTCTCCAGGCAAG	300	
Db	91042	AGAAAAAAT	CGACTT	CTACGCAAT	TCAACCAAT	TTCTGAGA	ATTTTCTCCAGGCAAG	91101	
QY	301	CGAT	TGCTTG	CGAGTAAGAAGAAAT	TGGTAAC	TATGACT	CGGATPCAGAGTAAGAGA	360	
Db	91102	CGAT	TGCTTG	CGAGTAAGAAGAAAT	TGGTAAC	TATGACT	CGGATPCAGAGTAAGAGA	91161	
QY	361	TTTTGAT	TTTTGTTAGTCT	CGATTCA	TCTCAT	TATGTGTTGA	TACAGAGATGACGAATCT	420	
Db	91162	TTTTGAT	TTTTGTTAGTCT	CGATTCA	TCTCAT	TATGTGTTGA	TACAGAGATGACGAATCT	91221	
QY	421	TTGAGGGT	TGACAGAGAGAA	AACTAAGA	AGATTTGG	GGATCAATCA	CAAAAACGATGGGGAAGA	480	
Db	91222	TTGAGGGT	TGACAGAGAGAA	AACTAAGA	AGATTTGG	GGATCAATCA	CAAAAACGATGGGGAAGA	91281	
QY	481	AAAGTCT	CAGGGTAA	AAATTTGGAAT	TAGAAAAACAT	GTGGAGAC	AGCTGACAGCTGGG	540	
Db	91282	AAAGTCT	CAGGGTAA	AAATTTGGAAT	TAGAAAAACAT	GTGGAGAC	AGCTGACAGCTGGG	91341	
QY	541	ATTATCT	GAGTTT	CACAGT	TAGACCTGT	GAGTTCA	TGCTTTTATCTGTTAACTTCGCA	600	
Db	91342	ATTATCT	GAGTTT	CACAGT	TAGACCTGT	GAGTTCA	TGCTTTTATCTGTTAACTTCGCA	91401	
QY	601	ACTCT	TTGATTAGT	TAACACAGAT	TAGAAAT	CTGAGAGCT	TTGGCGAGACGAGAGA	660	
Db	91402	ACTCT	TTGATTAGT	TAACACAGAT	TAGAAAT	CTGAGAGCT	TTGGCGAGACGAGAGA	91461	
QY	661	GAGA	GAAGAAAT	TGGCGGAGAGCT	GTACAT	TTCTCGAT	CTTACTTACGCATGTTATC	720	
Db	91462	GAGA	GAAGAAAT	TGGCGGAGAGCT	GTACAT	TTCTCGAT	CTTACTTACGCATGTTATC	91521	
QY	721	GCCT	TCGCGCTCT	GTGCCACT	TTTGGTCAT	TCCTGTAAG	GCATCATATTTCTGATTCG	780	
Db	91522	GCCT	TCGCGCTCT	GTGCCACT	TTTGGTCAT	TCCTGTAAG	GCATCATATTTCTGATTCG	91581	
QY	781	CAAG	CTTATTTCT	TAATTTCT	CAGATGAG	ACTGCA	CACTTTCTGGCTCCTTAATTCATA	840	
Db	91582	CAAG	CTTATTTCT	TAATTTCT	CAGATGAG	ACTGCA	CACTTTCTGGCTCCTTAATTCATA	91641	
QY	841	GCGAGA	AGTGCAT	AGCCGAT	ATCTTTAAAT	TCTTCCAT	TAAAGTAGAGCTGATPAAGG	900	
Db	91642	GCGAGA	AGTGCAT	AGCCGAT	ATCTTTAAAT	TCTTCCAT	TAAAGTAGAGCTGATPAAGG	91701	
QY	901	TCTTT	TATTTGAAAA	CCGATAT	CATATTTCT	CTCTG	CCCAATGAGATTTCCAGATTT	960	
Db	91702	TCTTT	TATTTGAAAA	CCGATAT	CATATTTCT	CTCTG	CCCAATGAGATTTCCAGATTT	91761	
QY	961	TTAG	GAACAAGG	TAGATTTAAAG	CTACAAAAT	CATATTTAG	TGAGTTTAAATATGC	1020	
Db	91762	TTAG	GAACAAGG	TAGATTTAAAG	CTACAAAAT	CATATTTAG	TGAGTTTAAATATGC	91821	
QY	1021	TAC	ACAGTTTCA	TCTTTAT	TATGATGAG	ATGACAAAT	GGTTTCAATGAGATGTTCTG	1080	
Db	91822	TAC	ACAGTTTCA	TCTTTAT	TATGATGAG	ATGACAAAT	GGTTTCAATGAGATGTTCTG	91881	

QY	1081	TTCTCAGACCCCTGGCTCTCTTTGGGGAAGTGGAGCTTCAGGACTTATCAACATATGCG	1140
Db	91882	TTACTCAGACCTTGGCTCTCTTTGGGGAAGTGGAGCTTCAAGACTTATCAACATATGCG	91941
QY	1141	CAAGACCCACTTGGATGCTGGAAAAATTTATATTTTGGTATGGAGCAATTTGAAGCTATTC	1200
Db	91942	CAAGACCCACTTGGATGCTGGAAAAATTTATATTTTGGTATGGAGCAATTTGAAGCTATTC	92001
QY	1201	TTTCAGCTGCTTGTGGCTGGTAAAGAAGTTGAGGCTCCAAATATCTCCACCCGGAACCGAC	1260
Db	92002	TTTCAGCTGCTTGTGGCTGGTAAAGAAGTTGAGGCTCCAAATATCTCCACCCGGAACCGAC	92061
QY	1261	CAGTTTACAGGTATGTTTCATTAATTAAGTCTGTGTGGTGAAGTTGAATGAGATGACAACTG	1320
Db	92062	CAGTTTACAGGTATGTTTCATTAATTAAGTCTGTGTGGTGAAGTTGAATGAGATGACAACTG	92121
QY	1321	TACGGTCTCTTTCTTGACAGGCCAATGCTGTGGCTCTTACTTTGTGACACTAGCACTA	1380
Db	92122	TACGGTCTCTTTCTTGACAGGCCAATGCTGTGGCTCTTACTTTGTGACACTAGCACTA	92181
QY	1381	TCCTGGCTCTTGGTGGTAAAGAATGTTCCGATCGACTTGGTGGTGAATGTTAATGTT	1440
Db	92182	TCCTGGCTCTTGGTGGTAAAGAATGTTCCGATCGACTTGGTGGTGAATGTTAATGTT	92241
QY	1441	TACGATCTTTAAGCTTGATTAAGTTGACCTATGTAAGTTAATTTCTGTGGCTAAT	1500
Db	92242	TACGATCTTTAAGCTTGATTAAGTTGACCTATGTAAGTTAATTTCTGTGGCTAAT	92301
QY	1501	CAGGTTGGAACCTTCACACCCCGCAATGCTATGATACACTGGGATGGAATTTTGGCG	1560
Db	92302	CAGGTTGGAACCTTCACACCCCGCAATGCTATGATACACTGGGATGGAATTTTGGCG	92361
QY	1561	ACTAATATTCGGAGAGCTTCATATTTTGTGTTTGTGTGACATAAAGTAAGTGTATTTG	1620
Db	92362	ACTAATATTCGGAGAGCTTCATATTTTGTGTTTGTGTGACATAAAGTAAGTGTATTTG	92421
QY	1621	CATGAATGCTGTCTTCTTTGTTTCTTAAGTCTCTATTTTGTAGAAATTTGACCTTTGTG	1680
Db	92422	CATGAATGCTGTCTTCTTTGTTTCTTAAGTCTCTATTTTGTAGAAATTTGACCTTTGTG	92481
QY	1681	TTCGCTAAACTTA-TTTTTTCTGTGACAGGGCCATGTTGCACCTTCATCACTAGTACTC	1739
Db	92482	TTCGCTAAACTTAATTTTTTCTGTGACAGGGCCATGTTGCACCTTCATCACTAGTACTC	92541
QY	1740	TGGTTCATGTGGTAACCTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1799
Db	92542	TGGTTCATGTGGTAACCTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	92601
QY	1800	-TCCTACTCTTTTGAAGAATGCTGTGGGATTTATTTAGGAACCTGGCTTACTAGC	1858
Db	92602	TTCCTACTCTTTTGAAGAATGCTGTGGGATTTATTTAGGAACCTGGCTTACTAGC	92661
QY	1859	CATCTCCCTAATAATTTTTGGCTCTTTCTCGACATTTGAGGTCAAGTTTCACTAAG	1918
Db	92662	CATCTCCCTAATAATTTTTGGCTCTTTCTCGACATTTGAGGTCAAGTTTCACTAAG	92721
QY	1919	TTTTCCAAATTTAGTAAACCTGTAAACGGAATGTTCTTAGAGTAAATTAATCTTTTATTTG	1978
Db	92722	TTTTCCAAATTTAGTAAACCTGTAAACGGAATGTTCTTAGAGTAAATTAATCTTTTATTTG	92781
QY	1979	ATTTAAGTACAGAAACACTAGTTTGGTGTGAATGTCTGTGGGTGAAGAAAGAACAAAG	2038
Db	92782	ATTTAAGTACAGAAACACTAGTTTGGTGTGAATGTCTGTGGGTGAAGAAAGAACAAAG	92841
QY	2039	TATGCAATGTTATTTAATCAATATCAAGTTGAGATCTTCTCAGCAATGTAATCTTTTG	2098
Db	92842	TATGCAATGTTATTTAATCAATATCAAGTTGAGATCTTCTCAGCAATGTAATCTTTTG	92901
QY	2099	CTCCAACTATCAATTAACAAATCCAAATTTGGGCAACTTTTCTTCACATGTAATTTTTCG	2158
Db	92902	CTCCAACTATCAATTAACAAATCCAAATTTGGGCAACTTTTCTTCACATGTAATTTTTCG	92961
QY	2159	CGTCGCTTTGCGAGGCGATGGAATTTGTACCTCGAATTTGTAGAGCTTTGACATCAAG	2218

Db 92962 CGTCGCTTTGGCAGGGGCGATGTGACCTCGAATTGGTAGACCTTGCATCAAG 93021
QY 2219 GGTGTTACTAATGAGATGGAATGATGCTGGGACAGTTCTTGGACAGTCACTGCG 2278
Db 93022 GTGTTACTAATGAGATGGAATGATGCTGGGACAGTTCTTGGACAGTCACTGCG 93081
QY 2279 ATAAACAGATATTTCTGATTAATTTCTATGTAATACTACTGACAGTCACTGCT 2338
Db 93082 ATAAACAGATATTTCTGATTAATTTCTATGTAATACTACTGACAGTCACTGCT 93141
QY 2339 GGTGTTCTTAAAGTCTTATTTTCATTTAGAGAGCTTACCTCTTCTGATTTGGATTC 2398
Db 93142 GGTGTTCTTAAAGTCTTATTTTCATTTAGAGAGCTTACCTCTTCTGATTTGGATTC 93201
QY 2399 CTGACTTTCTTTAACAGATGTAATTAATGGCAAGTATCTGATTCATTCAGTGGGAA 2458
Db 93202 CTGACTTTCTTTAACAGATGTAATTAATGGCAAGTATCTGATTCATTCAGTGGGAA 93261
QY 2459 CACCATCCTGATGCTGATGATGTACAAAATTTCTGCTGGGAGAGCTGTTATGGAA 2518
Db 93262 CACCATCCTGATGCTGATGATGTACAAAATTTCTGCTGGGAGAGCTGTTATTTGGAA 93321
QY 2519 CACCATGAGACTTGCATATGACGAGGATGCTTCTGATACATATATATAATTTATATCA 2578
Db 93322 CACCATGAGACTTGCATATGACGAGGATGCTTCTGATACATATATATAATTTATATCA 93381
QY 2579 TCACATCCTCATATAGGATGACAAAGAAAGAACGAGTAAATTAACAAAATTTCTCTCA 2638
Db 93382 TCACATCCTCATATAGGATGACAAAGAAAGAACGAGTAAATTAACAAAATTTCTCTCA 93441
QY 2639 CTGTTGGCTAATAATTTCTCTGAGCTTCTGTTATATGAGCTAAGGAGCAGCATATTAAT 2698
Db 93442 CTGTTGGCTAATAATTTCTCTGAGCTTCTGTTATATGAGCTAAGGAGCAGCATATTAAT 93501
QY 2699 TCACATATATGATCTGCTGCTTAAACATGACTTAAAGCTTACTTTTCAATGCTTTATTA 2758
Db 93502 TCACATATATGATCTGCTGCTTAAACATGACTTAAAGCTTACTTTTCAATGCTTTATTA 93561
QY 2759 TTTATGATCTTCCAGTGTGATCTATATATGCTGGGCTGCTAGTGGGCTGCTTCG 2818
Db 93562 TTTATGATCTTCCAGTGTGATCTATATATGCTGGGCTGCTAGTGGGCTGCTTCG 93621
QY 2819 TCTACACTTCTCCAGGATGTACCTTGTGAACCAACCCGCTGGAAGCTGGAAGTCAAGTGC 2878
Db 93622 TCTACACTTCTCCAGGATGTACCTTGTGAACCAACCCGCTGGAAGCTGGAAGTCAAGTGC 93681
QY 2879 ATATTTGATTTGCTTAAAGAAAGAAATCGCTTTGCTGTTAATCCATGAGAAAC 2938
Db 93682 ATATTTGATTTGCTTAAAGAAAGAAATCGCTTTGCTGTTAATCCATGAGAAAC 93741
QY 2939 TTTTATGATTTGCTTAAAGAAAGAAATCGCTTTGCTGTTAATCCATGAGAAAC 2998
Db 93742 TTTTATGATTTGCTTAAAGAAAGAAATCGCTTTGCTGTTAATCCATGAGAAAC 93801
QY 2999 TCCCAATTTTATAGAAATGACATTAATCTGTTATTTGAATTAATCTGATTTCTTCATTA 3058
Db 93802 TCCCAATTTTATAGAAATGACATTAATCTGTTATTTGAATTAATCTGATTTCTTCATTA 93861
QY 3059 C-AAACAGAGGGAGAGTAAGTGTATGATGATGATCAAAAGCTGCTGATATTAAG 3117
Db 93862 CAAACAGAGGGAGAGTAAGTGTATGATGATGATCAAAAGCTGCTGATATTAAG 93921
QY 3118 CTATGATTTGATTTGACAGATATCAATTTTCTAGTAGAGAGAAAGCTGTTCAATGA 3177
Db 93922 CTATGATTTGATTTGACAGATATCAATTTTCTAGTAGAGAGAAAGCTGTTCAATGA 93981
QY 3178 AACACCACTGATGCTACGATCATGCGGACA-AAATGCTATGCTCCGTAACCTAAAGCTTCA 3236
Db 93982 AACACCACTGATGCTACGATCATGCGGACA-AAATGCTATGCTCCGTAACCTAAAGCTTCA 94041
QY 3237 CTATATCGCTTTGGTGGATTTGGCAATATATCTTCGTTGCAAGATTTCTGTCAT 3296
|||||

Db 94042 CTATATCGCTTTGGTGGATTTGGCAATATATCTTCGTTGCAAGATTTCTGTCAT 94101
QY 3297 TTTCAATAAATCTATGACTGTGATATGACAAAGGACAGATTCAGAGAGACAACGGGAATG 3356
Db 94102 TTTCAATAAATCTATGACTGTGATATGACAAAGGACAGATTCAGAGAGACAACGGGAATG 94161
QY 3357 TTTGGTTGGGGAAGAGCCCGCAAGAGTGTATATTAAGAGAGAGAGAGAGAGAGAGAGAG 3416
Db 94162 TTTGGTTGGGGAAGAGCCCGCAAGAGTGTATATTAAGAGAGAGAGAGAGAGAGAGAGAG 94221
QY 3417 GAACCTTACGATTAATTTCTTTTACAGACCTTAACATTTTGTATGATTAATCTCTCTGCT 3476
Db 94222 GAACCTTACGATTAATTTCTTTTACAGACCTTAACATTTTGTATGATTAATCTCTCTGCT 94281
QY 3477 CTCTGCTTTGTTGCAAGATTTGGGCTGCTATATCAACATCTGTGTAACCTAAACCTA 3536
Db 94282 CTCTGCTTTGTTGCAAGATTTGGGCTGCTATATCAACATCTGTGTAACCTAAACCTA 94341
QY 3537 GTCCTCTTAAAGCTGTGAGAGTGTGAGACTATTTCTATATACAGTTTCAACCTAAATAA 3596
Db 94342 GTCCTCTTAAAGCTGTGAGAGTGTGAGACTATTTCTATATACAGTTTCAACCTAAATAA 94401
QY 3597 ATAGAGTTGTTATCATCATCAATGAGCTTTTGAATGTGAGAGTGGGAGATTTGCTGCTATTC 3656
Db 94402 ATAGAGTTGTTATCATCATCAATGAGCTTTTGAATGTGAGAGTGGGAGATTTGCTGCTATTC 94461
QY 3657 CATTAATGTTCCGAGATCTTAAAGTGTCTTCTGAGCCGCTACCGGCTCTCTGATTAAG 3716
Db 94462 CATTAATGTTCCGAGATCTTAAAGTGTCTTCTGAGCCGCTACCGGCTCTCTGATTAAG 94521
QY 3717 GTTAAATCTATTCACATCTTTTGTATTTGCAAGCTCTTCTGGAAGGAAAGCAGCT 3776
Db 94522 GTTAAATCTATTCACATCTTTTGTATTTGCAAGCTCTTCTGGAAGGAAAGCAGCT 94581
QY 3777 AACATTAATATCATCTTTTAAACAGATCTTGGACATCTTACGATATTTCTTC 3836
Db 94582 AACATTAATATCATCTTTTAAACAGATCTTGGACATCTTACGATATTTCTTC 94641
QY 3837 CCTCTCTCTTGTGATGACGCAAGAGAGAGAGATGACGAGTCCGATCAAAAGTAAAGCA 3896
Db 94642 CCTCTCTCTTGTGATGACGCAAGAGAGAGATGACGAGTCCGATCAAAAGTAAAGCA 94701
QY 3897 CAAATCTAATGATTAATTTTACATCTCTTTGATCTTATCTGATTTGCTGCAATTA 3956
Db 94702 CAAATCTAATGATTAATTTTACATCTCTTTGATCTTATCTGATTTGCTGCAATTA 94761
QY 3957 TTTGGGTTTGTGTTGTAAGAAAGTATGGAAGATTTGGAAGCTGATATTTGAG-AAAGT 4015
Db 94762 TTTGGGTTTGTGTTGTAAGAAAGTATGGAAGATTTGGAAGCTGATATTTGGAAG-AAAGT 94821
QY 4016 CAAATACAGATCATCTCCGGAATTTATGATTTGAAGAGTCTGTTCTCTCTATTTTC 4075
Db 94822 CAAATACAGATCATCTCCGGAATTTATGATTTGAAGAGTCTGTTCTCTCTATTTTC 94881
QY 4076 TACTTATTAAGTTAATTTGCAAGCTTGGATCATCTCAAAAGAGCCGCTCCAAACCAAAATG 4135
Db 94882 TACTTATTAAGTTAATTTGCAAGCTTGGATCATCTCAAAAGAGCCGCTCCAAACCAAAATG 94941
QY 4136 AAATGATGCGATACAGATCTTTTGTGATTTGATGCTATAGTTTCAATCTAG 4195
Db 94942 AAATGATGCGATACAGATCTTTTGTGATTTGATGCTATAGTTTCAATCTAG 95001
QY 4196 CTAGCTTATGATTTCCCTAGATTAATCAAGTATGAGCTGCTGCTTCTCAATTTATTAAG 4255
Db 95002 CTAGCTTATGATTTCCCTAGATTAATCAAGTATGAGCTGCTGCTTCTCAATTTATTAAG 95061
QY 4256 TTGATATTTGATTAATCTGCTTTCTTAATGATCAAAATAGTTGATTCATCATCTAATA 4315
Db 95062 TTGATATTTGATTAATCTGCTTTCTTAATGATCAAAATAGTTGATTCATCATCTAATA 95121
QY 4316 TTTT-AAATCATCAATGAAGCTTTTATTAATGATGAGGAGCTAATTTGAACCTTA 4374
Db 95122 TTTTAACTACATCAATGAAGCTTTTATTAATGATGAGGAGCTAATTTGAACCTTA 95181
|||||

AUTHORS Haas, B.J., Volkovsky, N., Town, C.D., Troukhan, M., Alexandrov, N., Feldmann, K.A., Flavell, R.B., White, O., and Salzberg, S.L.
TITLE Full-length messenger RNA sequences greatly improve genome annotation
JOURNAL Genome Biol. (2002) In press
REFERENCE 2 (bases 1 to 1194)
AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R., and Feldmann, K.
TITLE Full-length cDNA from Arabidopsis thaliana
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1194)
AUTHORS Brover, V., Troukhan, M., Alexandrov, N., Lu, Y.-P., Flavell, R., and Feldmann, K.
TITLE Direct Submission
JOURNAL Submitted (11-MAR-2002) Ceres, Inc., 3007 Malibu Canyon Road, Malibu, CA 90265, USA
COMMENT This clone sequence is one of 5,000 Ceres full-length cDNAs made available to TIGR and Genbank. The following quality assessment of this set was done by comparison with known proteins: two percent of the clones are estimated to be 5'-truncated; less than one percent are 3'-truncated; approximately two percent represent alternative splice variants, including unsplliced introns and spliced exons; one percent may contain premature stop codons; five percent may have frame shifts in a coding region. A sequence is considered to be 5'-truncated if it lacks the translation initiation start (ATG). A sequence is considered to be 3'-truncated if it lacks the C-terminal end of the encoded protein. Please note that these cDNA sequences are derived from the WS or Laer ecotypes and therefore may contain polymorphisms when compared to sequences from Col-0. Geneset carried out the library production and sequencing of the full-length clones. Ceres, Inc. carried out the clustering of the 5' sequences, selection of clones, and sequence assembly.

FEATURES

source 1. 1194
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="126074"
 229..981
CDS /codon_start=1
 /product="unknown"
 /protein_id="AA061619.1"
 /db_xref="GI:21537278"
 /translation="MOLVPSDDDDDKREQLDESTSSNEIIVAEKGRVVEGVSE
 LAETDDETLIVSGDPOPCRICDYGEDLLAPCKSGTKHRSKCDNNRSTKES
 FAFSHTCEKRAPEKLRANPADRWMLRLRPLLVAHAFIFISVOMIYAFGLIVYK
 FYGELREMGVEEHPHYRLAVLAIVGLIYGFIAIICGQKINENHYIVLAKOE
 LKREIVYERDCKNVPEDQSHVMEIKMGLY"

BASE COUNT

337 a 213 c 248 g 396 t

Query Match 4.9%; Score 237.4; DB 8; Length 1194;
Best Local Similarity 78.6%; Pred. No. 2.1e-36;
Matches 353; Conservative 0; Mismatches 1; Indels 95; Gaps 2;

QY 1 AACTTGTCTTCTCCACACAGATCTCTTCCGGCTGAGATTTCGTTAAAG 60
 |||||
DB 354 AACTTGTCTTCTCCACACAGATCTCTTCCGGCTGAGATTTCGTTAAAG 26
QY 61 ATGTTGATGTCGAATAAATCTGTTCTTATCATCATCATCATAGGACAT 120
 |||||
DB 295 ATGTTGATGTCGAATAAATCTGTTCTTATCATCATCATCATAGGACAT 236
QY 121 ATGTCATCTCTTAACATTAATAATTCTGGAATGAATAAATCAAGCAAA 180
 |||||
DB 235 ATGTCATCTCTTAACATTAATAATTCTGGAATGAATAAATCAAGCAAA 191
QY 181 GAATTCACACAGCATCAGAAACTAGATCTAGGTTCCGGGAGATTACCAAAACAT 240
 |||||
DB 190 ----- 191
QY 241 AGAAAAATGATCTTACACGAATTCACCAATTCGGGAAATTTCTTCCAGGCAAG 300
 |||||
DB 190 -----ACGAATTCACCAATTCGGGAAATTTCTTCCAGGCAAG 150

QY 301 CCATTGCTTCGAGTAAGAGAAATTTGTAACATCTGACTGCAATCAGAGTAAGAA 360
 |||||
DB 149 CCATTGCTTCGAGTAAGAGAAATTTGTAACATCTGACTGCAATCAGAGTAAGAA 90
QY 361 TTTGATTTTGTAGTCTCGATTCATCTATGTGTTGAATACAGAGATGAGAAATCT 420
 |||||
DB 89 TTTGATTTTGTAGTCTCGATTCATCTATGTGTTGAATACAGAGATGAGAAATCT 30
QY 421 TTGAGGCTTGAGAGAGAAAGCACTAAGAA 449
 |||||
DB 29 TTGAGGCTTGAGAGAGAAAGCACTAAGAA 1

RESULT 5
LOCUS A52625 1496 bp DNA linear PAT 12-DEC-1997
DEFINITION Sequence 1 from Patent EP0727489.
ACCESSION A52625
VERSION A52625.1 GI:2851857
KEYWORDS unclassified.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1496)
AUTHORS Chenivesse, X., Dupont, C., Lecain, E. and Pompon, D.
TITLE DNA sequence coding for an A. thaliana protein with delta-5,7-sterol delta-7-reductase, the protein, process for the production, transformed yeast strains and use
JOURNAL Patent: EP 0727489-A 1 21-AUG-1996;
COMMENT ROUSSEL DCLAF (FR)
 Other publication SK 18896 961001
 Other publication FR 2734839 961206
 Other publication JP 8289793 961105
 Other publication FI 960663 960816
 Other publication PL 312828 960819
 Other publication CA 2169524 960816
 Other publication FR 2730494 960814
 Other publication AU 4556796 960822.

FEATURES

source 1..1496
 /organism="unclassified"
 /db_xref="taxon:32644"
 76..1368
CDS /note="unnamed protein product"
 /codon_start=1
 /protein_id="CA03409.1"
 /db_xref="GI:2851858"
 /translation="MAETVHSPVITYASMLSLAFCPPEVLLMTYVHODGSYOTF
 GFEMNGVOGLINIPREPLTAMKILFCYGAFFAIIQLLLPGKREVPISPAQNPVY
 KANGIAAEVFLIATHGLIMWGCIFNPALVYDHGEISALIFGSEFCVLLYKGYA
 PSSDSGSCGMLIIDFYWGMELYTRICKSPDIKFTYTCRGRMSWALVATYICKOTE
 INGVKVSMDLVNTILMLVYTKFPEWMEGVNTMDIAHRAGEFYICWGLVWPSVYT
 SPGMVLNHPVELTOLAIVYTKFPEWMEGVNTMDIAHRAGEFYICWGLVWPSVYT
 IVASVYTSGETKTSILTSLSGMLARHFFHYPEILSAFETVVALPDNLAFVYVLT
 LLDPRAARDDDRCSKRYKXWKLXCEKVRRIIPGLY"

BASE COUNT

399 a 297 c 340 g 460 t

Query Match 4.6%; Score 224; DB 6; Length 1496;
Best Local Similarity 100.0%; Pred. No. 8.5e-34;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1049 ATGTACACAAATGTTATCAGAGATGTTCTGTACTCAGACCTTGGCTTTGGGA 1108
 |||||
DB 162 ATGTACACAAATGTTATCAGAGATGTTCTGTACTCAGACCTTGGCTTTGGGA 221
QY 1109 GATGAGTTCAAGGACTATACATATGCGACAGACCACTTGTATGTTGAAAT 1168
 |||||
DB 222 GATGAGTTCAAGGACTATACATATGCGACAGACCACTTGTATGTTGAAAT 281
QY 1169 TATATTTGCTATGAGCATTTGAAGCTATCTCAGCTGCTTCCCTGTAAGAGT 1228
 |||||

Db 282 TATATTTCCTATGAGCAATTTGAACTATTTTCAGCTGCTTCCTGTAAGAGCT 341

QY 1229 TGAGGGTCCAAATATCTCCAGCCGGAACCGACAGTTTACAGG 1272

Db 342 TGAGGGTCCAAATATCTCCAGCCGGAACCGACAGTTTACAGG 385

RESULT 6

AR079111

LOCUS AR079111 1496 bp DNA linear PAT 31-AUG-2000

DEFINITION Sequence 1 from patent US 5965417.

ACCESSION AR079111

VERSION AR079111.1 GI:10005857

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1496)

AUTHORS Chenivesse,X., Dupont,C., Lecain,E. and Pompon,D.

TITLE Arabidopsis thaliana proteins having

JOURNAL DELTA-5,7-sterol-DELTA-7-reductase activity

FEATURES

source Patent: US 5965417-A 1 12-OCT-1999;

Location/Qualifiers

1..1496

/organism="unknown"

BASE COUNT 399 a 297 c 340 g 460 t

ORIGIN

Query Match 4.6%; Score 224; DB 6; Length 1496;

Best Local Similarity 100.0%; Pred. No. 8.5e-34;

Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1049 ATGCTACACATATGCTTCATCAGATGCTTCTGTACTCAGACCTTGGCTTTGGGA 1108

Db 162 ATGCTACACATATGCTTCATCAGATGCTTCTGTACTCAGACCTTGGCTTTGGGA 221

QY 1109 GAATGAGTTCAGAGCACTTATCAACATATGAGCCAGACCTTGTATGCTTGAAGAT 1168

Db 222 GAATGAGTTCAGAGCACTTATCAACATATGAGCCAGACCTTGTATGCTTGAAGAT 281

QY 1169 TATATTTTGTATGAGCACTTATGAGCTATTTCTTCAGCTGCTTGTGCTGTAAGAGT 1228

Db 282 TATATTTTGTATGAGCACTTATGAGCTATTTCTTCAGCTGCTTGTGCTGTAAGAGT 341

QY 1229 TGAGGGTCCAAATATCTCCAGCCGGAACCGACAGTTTACAGG 1272

Db 342 TGAGGGTCCAAATATCTCCAGCCGGAACCGACAGTTTACAGG 385

RESULT 7

AR088377

LOCUS AR088377 1496 bp DNA linear PAT 07-SEP-2000

DEFINITION Sequence 1 from patent US 5989881.

ACCESSION AR088377

VERSION AR088377.1 GI:10015140

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1496)

AUTHORS Chenivesse,X., Dupont,C., Lecain,E. and Pompon,D.

TITLE Nucleic acid molecules encoding delta-5,7-sterol, delta-7 reductase

JOURNAL Patent: US 5989881-A 1 23-NOV-1999;

FEATURES

source Patent: US 5989881-A 1 23-NOV-1999;

Location/Qualifiers

1..1496

/organism="unknown"

BASE COUNT 399 a 297 c 340 g 460 t

ORIGIN

Query Match 4.6%; Score 224; DB 6; Length 1496;

Best Local Similarity 100.0%; Pred. No. 8.5e-34;

Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1049 ATGCTACACATATGCTTCATCAGATGCTTCTGTACTCAGACCTTGGCTTTGGGA 1108

Db 162 ATGCTACACATATGCTTCATCAGATGCTTCTGTACTCAGACCTTGGCTTTGGGA 221

QY 1109 GAATGAGTTCAGAGCACTTATCAACATATGAGCCAGACCTTGTATGCTTGAAGAT 1168

Db 222 GAATGAGTTCAGAGCACTTATCAACATATGAGCCAGACCTTGTATGCTTGAAGAT 281

QY 1169 TATATTTTGTATGAGCACTTATGAGCTATTTCTTCAGCTGCTTGTGCTGTAAGAGT 1228

Db 282 TATATTTTGTATGAGCACTTATGAGCTATTTCTTCAGCTGCTTGTGCTGTAAGAGT 341

QY 1229 TGAGGGTCCAAATATCTCCAGCCGGAACCGACAGTTTACAGG 1272

Db 342 TGAGGGTCCAAATATCTCCAGCCGGAACCGACAGTTTACAGG 385

RESULT 8

ATU49398

LOCUS ATU49398 1497 bp mRNA linear PLN 16-MAY-1996

DEFINITION Arabidopsis thaliana sterol delta-7 reductase (ST7R) mRNA, complete cds.

ACCESSION U49398

VERSION U49398.1 GI:1245181

KEYWORDS

SOURCE Arabidopsis thaliana.

ORGANISM Arabidopsis thaliana.

REFERENCE 1 (bases 1 to 1497)

AUTHORS Lecain,E., Chenivesse,X., Spagnoli,R. and Pompon,D.

TITLE Cloning by metabolic interference in yeast and enzymatic characterization of Arabidopsis thaliana sterol delta-7 reductase

JOURNAL J. Biol. Chem. 271 (18), 10866-10873 (1996)

REFERENCE 2 (bases 1 to 1497)

AUTHORS Pompon,D., Lecain,E., Chenivesse,X. and Spagnoli,R.

TITLE Direct Submission

JOURNAL Submitted (17-FEB-1996) Denis Pompon, Centre de genetique Moleculaire, CNRS, Avenue de la Terrasse, Gif-sur-Yvette 91190, France

FEATURES

source Location/Qualifiers

1..1497

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/tissue-type="full seedling"

1..1497

/gene="ST7R"

76..1368

/gene="ST7R"

/note="NADPH dependent; sterol reductase"

/codon_start=1

/product="sterol delta-7 reductase"

/protein_id="AAC49278.1"

/db_xref="GI:1245182"

/translation="MAETVSPVITVYASMLSLAFCPPIVILMTVMHODGSVTQTE GFPEWNGVQGLINWPRPLDIANKIIFCYGAFAIILQLLPGRVGGPISPAGNRPVY KANGLAAYEVTLATHGLMWEFGJFNPAIVYDHGELFSAIFGSEFCVLYLTKGYA PSSSDSGSCNLIIDPFYGMELIIPRIGKSDIDVFNCRRGMSMAVLATVCIKOYE INKRVSDSMIVNTIIMLVYTKTFPMWAGVWNTMDIAHBRAGFYICMGILVWPSYTT SPQWLVNHPVEIGTDLATILVAGILCTIKYDCRQROEERRINGKLVWGRARSK IVASVTTTSGEFTSLTSGWGLARHFPVPELISAFETVVPALDELFLAVFYLT ILLFDRAKRDDDCRSKRYGKRWLYCEKVKYRIIPGIV"

BASE COUNT 400 a 297 c 340 g 460 t

ORIGIN

Query Match 4.6%; Score 224; DB 8; Length 1497;

Best Local Similarity 100.0%; Pred. No. 8.5e-34;

Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1049 ATGGTACAAATGTTTCATCAGATGTTCTGTACTACAGACCTTGGCTTTTGGGA 1108
 |||||
 Db 162 ATGTACAAATGTTTCATCAGATGTTCTGTACTACAGACCTTGGCTTTTGGGA 221
 |||||
 QY 1109 GAATGAGTTTCAGAGCTTATCAACATATGCGCAAGCCACCTTTGATTTCTGGGAAAT 1168
 |||||
 Db 222 GAATGAGTTTCAGAGCTTATCAACATATGCGCAAGCCACCTTTGATTTCTGGGAAAT 281
 |||||
 QY 1169 TATATTTGCTATGAGACATTTGAAGCTATTTTCAAGCTCTTCTGCTGTAAGAGAT 1228
 |||||
 Db 282 TATATTTGCTATGAGACATTTGAAGCTATTTTCAAGCTCTTCTGCTGTAAGAGAT 341
 |||||
 QY 1229 TGAGGTCACATATCTTCACCCGAAACCGACCATTTTACAAAG 1272
 |||||
 Db 342 TGAGGTCACATATCTTCACCCGAAACCGACCATTTTACAAAG 385
 |||||

RESULT 9
 AP004850 125348 bp DNA linear HTG 23-MAR-2002
 LOCUS Oryza sativa (japonica cultivar-group) chromosome 2 clone
 DEFINITION OJ1342.D02, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
 ACCESSION AP004850.1 GI:19698270
 VERSION HTG; HTGS_PHASE2.
 KEYWORDS Oryza sativa (japonica cultivar-group) (cultivar:Niponbare) DNA,
 SOURCE clone:OJ1342.D02.
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
 TITLE Oryza sativa niponbare(GA3) genomic DNA, chromosome 2, BAC
 JOURNAL clone:OJ1342.D02
 PUBLISHED ONLY in Database (2002)
 2 (bases 1 to 125348)
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail:tsasakikenia.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 COMMENT The nucleotide sequence of this BAC clone was generated by
 combining Monsanto and RGP-Japan sequencing data.
 NOTE: It currently consists of 1 contigs. Gaps between the contigs
 are represented as runs of N. The order of the pieces is believed
 to be correct as given, however the sizes of the gaps between them
 are based on estimates that have provided by the submitter. This
 sequence will be replaced by the finished sequence as soon as it is
 available and the accession number will be preserved.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 FEATURES
 source
 1. 125348
 Location/Qualifiers
 /organism="Oryza sativa (japonica cultivar-group)"
 /cultivar="Niponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="OJ1342.D02"
 BASE COUNT 35739 a 27515 c 26919 g 35124 t 51 others.
 ORIGIN

Query Match 2.8%; Score 135.6; DB 2; Length 125348;
 Best Local Similarity 55.2%; Pred. No. 9.9e-17;
 Matches 403; Conservative 0; Mismatches 249; Indels 78; Gaps 4;

QY 3209 TATGCCATATCCGTACTACACTATATCCGCTTTGGTGGCACTTGGCAATAT 3268
 |||||
 Db 124567 TGTGATTTGTTCTAAACAGTTGAATGAATATCTGTTAAATTTGTCAGACTGACCTCT 124626
 |||||

QY 3269 ACATTCCTGTCAGGAATTTCTGTCATTTACATTAACATGACCTGTATGACAAAGCC 3328
 |||||
 Db 124627 CAATTCCTCTTCTGATATATTTGATATATATATATGATGATGATGTCAGCCGC 124686
 |||||
 QY 3329 AAGAGTTTCAGAGACAAACCGGAATGTTGGTGGGGAAGAGCCCGTCAAGAGTGT 3388
 |||||
 Db 124687 AAGAAATTCGCGGGAACAAATGGAATGTCATATATGSGGCAACCTCATCTAAGTTG 124746
 |||||
 QY 3389 GATCATATATAGCCCAACTAGCTTAGTGAATGCTATATATCTTTTACAGACTTAAC 3448
 |||||
 Db 124747 AAGCATTTAA-----TATTTAGCTGCTTCCAAATTAACAAATAGCA 124787
 |||||
 QY 3449 ATTTTGATGTATGATGATCTCTTGGTCTGCTGCTGCTTGGTGGCACTTGGCGCTGT 3508
 |||||
 Db 124788 TGTTTATACATTTGCTGACGCTTTTATTTTGGTGAACCTT--AGATTTGCTTCCAT 124845
 |||||
 QY 3509 ACTACACATCTGATGAAACCTAAACATGATCTTCTTACGCTGATGATGATGAT 3568
 |||||
 Db 124846 CAGACTACAAATGAGAAACAAACAAAGAGCTTCTTCTGACTTCTGATGATATATAT 124905
 |||||
 QY 3569 TCTACATTTACAGTTTACACCATATAAATAGAGCTTGTATCATCATAGCTTTGAA- 3627
 |||||
 Db 124906 TCGCATTTCTTTTACATAGTTCATCAAAAGATTTGATACATGTTTCTTCAATAG 124965
 |||||
 QY 3628 -----TGTGCAAGGTGGGATTTGG 3645
 |||||
 Db 124966 AGCCATGATGATGCTTTTGGAGCATGACCCCTCATTTTCCATTTTGGAGGCGCTTGT 125025
 |||||
 QY 3646 CTCGTCATTTCCATATATGTTCTCTGAGATCTTAAGTCTTCTTCTGAGACCTTACGCGCTC 3705
 |||||
 Db 125026 CTCGTCATTTCCATATATGTTCTCTGAGATCTTAAGTCTTCTTCTGAGACCTTACGCGCTC 125085
 |||||
 QY 3706 TCTTGATATAGCTAAATATATATATATATATTTGCTATATTTGGAAGCTTCTTGGCAAGT 3765
 |||||
 Db 125086 TTTTGTATCATATGTA-----GCACACCTGATATACATTAATATCTGACATTT 125131
 |||||
 QY 3766 GAACGACATTAACATTAATATATATATATATATTTTAAACGATCTTGGATATCTTATAGT 3825
 |||||
 Db 125132 ACCCAGATGAT 125191
 |||||
 QY 3826 CATATTTCTCACCCCTTCTCTCTTTGATGACGCAAGAGAGATGACGATGCGCATC 3885
 |||||
 Db 125192 GATCTTTTGACCATATATGCTTTGACGACGCTAAAGGATGATGACGATGCTATC 125251
 |||||
 QY 3886 AAAGTATAGCA 3895
 |||||
 Db 125252 AAAGTATAGTA 125261
 |||||

RESULT 10
 AP005514 178138 bp DNA linear HTG 11-JUL-2002
 LOCUS Oryza sativa (japonica cultivar-group) chromosome 2 clone
 DEFINITION OSJNBa0035A24, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
 ACCESSION AP005514
 VERSION HTG; HTGS_PHASE2.
 KEYWORDS Oryza sativa (japonica cultivar-group) (cultivar:Niponbare) DNA,
 SOURCE clone:OSJNBa0035A24.
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE 1
 AUTHORS Sasaki,T., Matsumoto,T. and Katayose,Y.
 TITLE Oryza sativa niponbare(GA3) genomic DNA, chromosome 2, BAC
 JOURNAL clone:OSJNBa0035A24
 PUBLISHED ONLY in Database (2002)
 2 (bases 1 to 178138)
 AUTHORS Sasaki,T., Matsumoto,T. and Katayose,Y.
 TITLE Direct Submission
 JOURNAL Submitted (10-JUL-2002) Takuji Sasaki, National Institute of

Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsukuba@affrc.go.jp, URL: http://rpp.dna.affrc.go.jp/
Tel: 81-298-38-7441, Fax: 81-298-38-7468)

NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

Source

1. 178138
/organism="Oryza sativa (japonica cultivar-group)"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OSJNB0035A24"

BASE COUNT

51522 a 37116 c 36733 g 52716 t 51 others

ORIGIN

Query Match

2.8%; Score 135.6; DB 2; Length 178138;

Best Local Similarity 55.2%; Pred. No. 9,4e-17;
Matches 403; Conservative 0; Mismatches 249; Indels 78; Gaps 4;

```

QY 3209 TATGCTATCTCCGACCTNAGACTTCACTATATCCGCTTTGGTTGCGAGTTGCCAATAT 3268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 24918 TGTGATGTTGTTCTAACAGTTGAATGAAATCTGTTAAATGTTGACGACGACCTCT 24977
QY 3269 ACATTCCTGTTGAGAGATCTGTGATTTACATAACTAATAGTATGATAGCAAAAGCC 3328
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 24978 CAATTCCTCTGCTGGAATATTTGCAATATATATACATATGACTGTGATCTGACGCC 25037
QY 3329 AAGAGTTGAGAGAGACAAACGGGAAATGTTGGTTGGGAGAGAGCCCGCTCAAGTGT 3388
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25038 AAGAAATCCGTCGACAAATATGGGAATGTCATATATGGGCAAAAGCTCATTAAGTTG 25097
QY 3389 GATCATATATAGGCCCACTAGCTTATGATCTATCTATCTCTTTCAACAGCTAATC 3448
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25098 AAGCAATTAATATATATATATATATATATATATATATATATATATATATATATAT 25138
QY 3449 ATTTGATGATGATGATCTCTCTGCTGCTCGGTTTGTGAGATTTGGCGTGTAT 3508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25139 TGTATATACATCTGACGCTCTTTATTTGGTGAACCTT--AGATTTGCTGCTCTAT 25136
QY 3509 ACTACAACATCTGTGAATTAACCTAATCTCTCTTAACGCTGATGATGAGACTAT 3568
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25197 CAGACTACAAATATGAGAAACAAACAGACGCTCTCTGATCTGATGATGATATATAT 25256
QY 3569 TCTACATATATACAGTAAACCAATTAACAGCTTGTTCATCATCATGCGCTTTGAA- 3627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25257 TCGTCATTTCTCTTACATATGATCAAAAGTATTTGACCATGTTTGTTCATATG 25316
QY 3628 -----TGTGACAGTGGGAGATTTG 3645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25317 AGGCATGATGCTTTTGAATGACATCCCTCATTTTCCATTTGGACAGTGGGCTTGT 25376
QY 3646 CTTGCTATTTCCATATATGTTCTGAGATCTTAAGTCTTTCTTCTGACCGATACCGCTC 3705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25377 CTTGCTATTTCCATATATGTTCTGAGATCTTAAGTCTTTTCTTCTGACAGTTCAGCTC 25436
QY 3706 TCTCATATACGTAATAATATCATCATTTGCTATATTCGAAGCTCTTCTTGAAGT 3765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25437 TTTTATATCATGTA-----GCACACCTCATTAACAATTTATCTCGACTTT 25482
QY 3766 GAAAGCACTTATCATATTAATCATCTTTGTTTAAACAGCTTTTGGCATCATCTTACCT 3825
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 25483 AGCAGAGATGTAATATACATCTGCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 25542
QY 3826 CATATTTCTCACCTCTCTCTTTGATGAGAGCAAGAGAGAGATGACCATGCGCATC 3885

```

Db 25543 GATCTTCTGACCAATATGCTTTCGACCGACGTAAGGAGATGATGACCGATGCTATC 25602

QY 3886 AAGTAAGCA 3895

Db 25603 AAGTAAGCA 25612

RESULT 11

AC013349

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

NOTE: This record contains 150 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 930: contig of 930 bp in length
* 931 1030: gap of 100 bp
* 1031 1955: contig of 925 bp in length
* 1956 2055: gap of 100 bp
* 2056 2959: contig of 904 bp in length
* 2960 3059: gap of 100 bp
* 3060 4006: contig of 947 bp in length
* 4007 4106: gap of 100 bp

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: W1BR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: LA134
Center clone name: 22_K_1

Submitted (06-NOV-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced g1:6910730.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center

40470	40569:	gap of	100 bp	
40570	41446:	contig of 877 bp	In Length	
41447	41546:	gap of 100 bp		
41547	42418:	contig of 872 bp	In Length	
42419	42518:	gap of 100 bp		
42519	43377:	contig of 859 bp	In Length	
43378	43477:	gap of 100 bp		
43478	4433:	contig of 856 bp	In Length	
44334	44433:	gap of 100 bp		
44334	45324:	contig of 891 bp	In Length	
45325	45424:	gap of 100 bp		
45425	46337:	contig of 913 bp	In Length	
46338	46437:	gap of 100 bp		
46338	47340:	contig of 903 bp	In Length	
47341	47341:	gap of 100 bp		
47341	48330:	contig of 890 bp	In Length	
48331	48430:	gap of 100 bp		
48431	49328:	contig of 898 bp	In Length	
49329	49428:	gap of 100 bp		
49429	50312:	contig of 884 bp	In Length	
50313	50412:	gap of 100 bp		
50413	51297:	contig of 880 bp	In Length	
51293	51392:	gap of 100 bp		
51393	52229:	contig of 837 bp	In Length	
52230	52329:	gap of 100 bp		
52330	53166:	contig of 857 bp	In Length	
53187	53286:	gap of 100 bp		
53387	54213:	contig of 933 bp	In Length	
54220	54319:	gap of 100 bp		
54320	55206:	contig of 887 bp	In Length	
55207	55306:	gap of 100 bp		
55307	56208:	contig of 902 bp	In Length	
56209	56308:	gap of 100 bp		
56309	57218:	contig of 910 bp	In Length	
57219	57318:	gap of 100 bp		
57319	58231:	contig of 913 bp	In Length	
58232	58331:	gap of 100 bp		
58332	59246:	contig of 915 bp	In Length	
59247	59346:	gap of 100 bp		
59347	60241:	contig of 895 bp	In Length	
60242	60341:	gap of 100 bp		
60342	61250:	contig of 909 bp	In Length	
61251	61350:	gap of 100 bp		
61351	62216:	contig of 866 bp	In Length	
62217	62316:	gap of 100 bp		
62317	63190:	contig of 874 bp	In Length	
63191	63290:	gap of 100 bp		
63291	64198:	contig of 908 bp	In Length	
64199	64298:	gap of 100 bp		
64299	65187:	contig of 889 bp	In Length	
65188	65287:	gap of 100 bp		
65288	66177:	contig of 890 bp	In Length	
66178	66277:	gap of 100 bp		
66278	67188:	contig of 911 bp	In Length	
67189	67288:	gap of 100 bp		
67289	68193:	contig of 905 bp	In Length	
68194	68293:	gap of 100 bp		
68294	69212:	contig of 919 bp	In Length	
69213	69312:	gap of 100 bp		
69313	70204:	contig of 892 bp	In Length	
70205	70304:	gap of 100 bp		
70305	71214:	contig of 910 bp	In Length	
71215	71314:	gap of 100 bp		
71315	72222:	contig of 908 bp	In Length	

```
Oy 1379 TACCTTGGCTTTGGTGGTAGAAGATGTTTTCCGATGACTTTGGTTGGTTTAAGTTATTG 1438
      | | | | | | | | | | | | | | | | | |
Db 3425 TTTTNTTTTTTTTTTTTTNNNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3486
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 15:16:02 ; Search time 194 Seconds
(without alignments)
11213.358 Million cell updates/sec

Title: US-09-817-774-30

Perfect score: 4879
Sequence: 1 aactgtctctctccacaa.....atctaaacaatttactg 4879

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues

Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications -NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCRUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	4879	100.0	4879	10	US-09-817-774-30
2	4868	99.8	4880	10	US-09-817-774-28
3	4866.4	99.7	4880	10	US-09-817-774-32
4	4866.4	99.7	4880	10	US-09-817-774-34
5	4866.4	99.7	4880	10	US-09-817-774-36
6	4866.4	99.7	4880	10	US-09-817-774-38
7	4866.4	99.7	4880	10	US-09-817-774-40
8	4866.4	99.7	4880	10	US-09-817-774-42
9	541.4	11.1	543	9	US-09-938-842A-4796
10	267.6	5.5	2000	9	US-09-938-842A-3847
11	249.8	5.1	1531	10	US-09-817-774-44
12	224	4.6	1299	9	US-09-938-842A-2112
13	115	2.4	273	10	US-09-878-574-15028
14	58.8	1.2	529	10	US-09-983-965-2109
15	55.6	1.1	446	10	US-09-960-352-3400
16	53.2	1.1	1423	9	US-10-001-857-100
17	52.2	1.1	516	10	US-09-960-352-5785
18	50.2	1.0	480	10	US-09-960-352-5301
19	48.8	1.0	337	10	US-09-960-352-6976

20	48.8	1.0	344	10	US-09-960-352-1036	Sequence 1036, Ap
21	48	1.0	1267	12	US-10-001-843-45	Sequence 45, Ap1
22	47.8	1.0	442	10	US-09-960-352-12911	Sequence 12911, A
23	47.6	1.0	425	10	US-09-960-352-4010	Sequence 4010, Ap
24	47.4	1.0	341	10	US-09-960-352-12302	Sequence 12302, A
25	47.4	1.0	408	10	US-09-960-352-12301	Sequence 12301, A
26	47.4	1.0	428	10	US-09-960-352-573	Sequence 573, App
27	46.8	1.0	335913	9	US-09-754-853A-2	Sequence 2, Ap11
28	46.8	1.0	335913	9	US-09-754-853A-3	Sequence 3, Ap11
29	45.8	0.9	430	10	US-09-783-590-7886	Sequence 7886, Ap
30	45.6	0.9	390	10	US-09-960-352-3640	Sequence 3640, Ap
31	45.4	0.9	491	10	US-09-878-574-4301	Sequence 4301, Ap
32	45.4	0.9	1428	9	US-10-001-857-79	Sequence 79, Ap1
33	45.2	0.9	283	10	US-09-960-352-9095	Sequence 9095, Ap
34	44.8	0.9	431	10	US-09-960-352-5558	Sequence 5558, Ap
35	44.6	0.9	418	10	US-09-960-352-3879	Sequence 3879, Ap
36	44.6	0.9	419	10	US-09-960-352-11234	Sequence 11234, A
37	44.6	0.9	452	10	US-09-960-352-9806	Sequence 9806, Ap
38	44.4	0.9	393	10	US-09-960-352-4582	Sequence 4582, Ap
39	44.2	0.9	463	10	US-09-960-352-7186	Sequence 7186, Ap
40	44	0.9	290	10	US-09-294-093B-6170	Sequence 6170, Ap
41	44	0.9	397	10	US-09-960-352-13784	Sequence 13784, A
42	44	0.9	4167	10	US-09-764-878-282	Sequence 282, App
43	44	0.9	4167	10	US-09-764-860-1145	Sequence 1145, Ap
44	44	0.9	4167	10	US-09-764-846-345	Sequence 345, App
45	43.8	0.9	545	10	US-09-878-574-4239	Sequence 4239, Ap

ALIGNMENTS

RESULT 1
US-09-817-774-30
Sequence 30, Application US/09817774
Patent No. US2002012011A1
GENERAL INFORMATION:
APPLICANT: CHOE, Sunghwa
APPLICANT: FELDMANN A., Kenneth
TITLE OF INVENTION: Dwf5 MUTANTS
FILE REFERENCE: 2225-0020 / 91020.002
CURRENT FILING DATE: 2001-03-26
PRIORITY APPLICATION NUMBER: 60/192,202
PRIOR FILING DATE: 2000-03-27
NUMBER OF SEQ ID NOS: 45
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 30
LENGTH: 4879
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: CDS
LOCATION: (671)..(757)
OTHER INFORMATION: dwf5-1
NAME/KEY: CDS
LOCATION: (1050)..(1271)
OTHER INFORMATION: dwf5-1
NAME/KEY: CDS
LOCATION: (1340)..(1396)
OTHER INFORMATION: dwf5-1
NAME/KEY: CDS
LOCATION: (1505)..(1606)
OTHER INFORMATION: dwf5-1
NAME/KEY: CDS
LOCATION: (1711)..(1776)
OTHER INFORMATION: dwf5-1
NAME/KEY: CDS
LOCATION: (2174)..(2287)
OTHER INFORMATION: dwf5-1
NAME/KEY: CDS
LOCATION: (2418)..(2546)
OTHER INFORMATION: dwf5-1
NAME/KEY: CDS

LOCATION: (2776)..(2874)
 OTHER INFORMATION: dwf5-1
 NAME/KEY: CDS
 LOCATION: (3259)..(3384)
 OTHER INFORMATION: dwf5-1
 NAME/KEY: CDS
 LOCATION: (3494)..(3559)
 OTHER INFORMATION: dwf5-1
 NAME/KEY: CDS
 LOCATION: (3636)..(3716)
 OTHER INFORMATION: dwf5-1
 NAME/KEY: CDS
 LOCATION: (3806)..(3889)
 OTHER INFORMATION: dwf5-1
 NAME/KEY: CDS
 LOCATION: (3982)..(4179)
 OTHER INFORMATION: dwf5-1
 US-09-817-774-30

Query Match 100.0%; Score 4879; DB 10; Length 4879;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4879; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTGTCCTTCCTCCACAACAGATCTCTCTTCGGCTCTACGATTTGTTTAAAG 60
 DB 1 AACTGTCCTTCCTCCACAACAGATCTCTCTTCGGCTCTACGATTTGTTTAAAG 60
 QY 61 AAGTATTCGTCGATTAATCTGTTCTCTCTATCATCATCATCATGAGGACACTA 120
 DB 61 AAGTATTCGTCGATTAATCTGTTCTCTCTATCATCATCATCATGAGGACACTA 120
 QY 121 ATTGCATCTCTTAACATCTAAATTTCTGGAATGAAAAAACTGAGAACCCAAA 180
 DB 121 ATTGCATCTCTTAACATCTAAATTTCTGGAATGAAAAAACTGAGAACCCAAA 180
 QY 181 GAATTAACACAGCATCAAAAACTAGATCTAGTTTCCCGGAGTTACCAAAACAT 240
 DB 181 GAATTAACACAGCATCAAAAACTAGATCTAGTTTCCCGGAGTTACCAAAACAT 240
 QY 241 AGAAAAATCGATCTACTTACAGCAATTCGCAAAATTTCTCCAGAGCAAG 300
 DB 241 AGAAAAATCGATCTACTTACAGCAATTCGCAAAATTTCTCCAGAGCAAG 300
 QY 301 CCATTCCTTCGAGTAAGAAATTTGTAACATATGACTGCGAATCAGAGTGAAGAA 360
 DB 301 CCATTCCTTCGAGTAAGAAATTTGTAACATATGACTGCGAATCAGAGTGAAGAA 360
 QY 361 TTTGATTTTGTAGTCTCATCTATCATCTATGTTGAATACAGAGTACAGAAATCT 420
 DB 361 TTTGATTTTGTAGTCTCATCTATCATCTATGTTGAATACAGAGTACAGAAATCT 420
 QY 421 TTGAGGTTGAG 480
 DB 421 TTGAGGTTGAG 480
 QY 481 AAAGTCTCAAGGGTAAATTTGGAATTAAGAAAAACATGTGAGAGACAGCTGAG 540
 DB 481 AAAGTCTCAAGGGTAAATTTGGAATTAAGAAAAACATGTGAGAGACAGCTGAG 540
 QY 541 ATTATCTAGTTTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 600
 DB 541 ATTATCTAGTTTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 600
 QY 601 ACTCTGTTAGTATGTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 DB 601 ACTCTGTTAGTATGTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 QY 661 GAAG 720
 DB 661 GAAG 720
 QY 721 GCTTCGCTTCCTGTCACCTTTGTCATTTCTCTGTAAGCTCAATTTCTGATTCG 780
 DB 721 GCTTCGCTTCCTGTCACCTTTGTCATTTCTCTGTAAGCTCAATTTCTGATTCG 780

DB 721 GCTTCGCTTCCTGTCACCTTTGTCATTTCTCTGTAAGCTCAATTTCTGATTCG 780
 QY 781 CAAGTCCTTTATTTAGTTTCTAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 840
 DB 781 CAAGTCCTTTATTTAGTTTCTAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 840
 QY 841 GCGAGAGAGTGCATAGCCGATATCTTTAAATCTTTCCCATTAAGTAGAGAGAGAGAG 900
 DB 841 GCGAGAGTGCATAGCCGATATCTTTAAATCTTTCCCATTAAGTAGAGAGAGAGAGAG 900
 QY 901 TCTTTAATGAAAAAGGATTCATATGATTTCTCTCTGCGCAATGTAGTTCCAGATT 960
 DB 901 TCTTTAATGAAAAAGGATTCATATGATTTCTCTCTGCGCAATGTAGTTCCAGATT 960
 QY 961 TTAGGAACCAAGGTTAGATTTAAAGCTACAAATCATATTTAGTGGTTTAAGTATGC 1020
 DB 961 TTAGGAACCAAGGTTAGATTTAAAGCTACAAATCATATTTAGTGGTTTAAGTATGC 1020
 QY 1021 TCACAGTTTCATTTCTTTATGATGAGATGTACACAAATGTTTCATCAGATGTTCTG 1080
 DB 1021 TCACAGTTTCATTTCTTTATGATGAGATGTACACAAATGTTTCATCAGATGTTCTG 1080
 QY 1081 TTACTCAGACCTTTGGCTTTCTTTGGGAGATGAGATTCAAGACTTATCAATATGCC 1140
 DB 1081 TTACTCAGACCTTTGGCTTTCTTTGGGAGATGAGATTCAAGACTTATCAATATGCC 1140
 QY 1141 CAAGCCCACTTTGATTTGTTGAAAAATTAATTTGCTATGAGAGATTTGAAGCTATTTC 1200
 DB 1141 CAAGCCCACTTTGATTTGTTGAAAAATTAATTTGCTATGAGAGATTTGAAGCTATTTC 1200
 QY 1201 TTCAGTCTCTTCCTCCGTAAGAAAGTTGAGGTCATATCTCCAGCCGAACCGAC 1260
 DB 1201 TTCAGTCTCTTCCTCCGTAAGAAAGTTGAGGTCATATCTCCAGCCGAACCGAC 1260
 QY 1261 CAGTTTACAGATTTGTTCAATTTAGTCTGTGTGTGTGAGAGAGATGACATACAACTG 1320
 DB 1261 CAGTTTACAGATTTGTTCAATTTAGTCTGTGTGTGTGAGAGAGATGACATACAACTG 1320
 QY 1321 TACGTTCTTTCTTGGAGGCAATGCTGCGCTTACTTTGTGACATAGCAACTTA 1380
 DB 1321 TACGTTCTTTCTTGGAGGCAATGCTGCGCTTACTTTGTGACATAGCAACTTA 1380
 QY 1381 TCTTGTCTTTGTTGTTGTAAGAAATGTTTCCGATCTTTGTTGTTAAGTTATTT 1440
 DB 1381 TCTTGTCTTTGTTGTTGTAAGAAATGTTTCCGATCTTTGTTGTTAAGTTATTT 1440
 QY 1441 TACGATCTTAAGCTTGAATTAAGTGTGCACTATGTGATTAATTTCTGCGTATTT 1500
 DB 1441 TACGATCTTAAGCTTGAATTAAGTGTGCACTATGTGATTAATTTCTGCGTATTT 1500
 QY 1501 CAGTTTGAATCTTCAACCTCGCAATTTGATGATCACTTGGTGAATATTTTCGCG 1560
 DB 1501 CAGTTTGAATCTTCAACCTCGCAATTTGATGATCACTTGGTGAATATTTTCGCG 1560
 QY 1561 ACTAATTTGGAAGCTTCATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1620
 DB 1561 ACTAATTTGGAAGCTTCATATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1620
 QY 1621 CATGAATGTGTGTTCTTTGTTCTTGTAGTCTCATATTTTGTGAATTAACCTTTG 1680
 DB 1621 CATGAATGTGTGTTCTTTGTTCTTGTAGTCTCATATTTTGTGAATTAACCTTTG 1680
 QY 1681 TTGCTAAACCTTATTTTCTTGTGACAGGCAATTTGACCTTCATCAATGACTCT 1740
 DB 1681 TTGCTAAACCTTATTTTCTTGTGACAGGCAATTTGACCTTCATCAATGACTCT 1740
 QY 1741 GGTTCATGTGTAACTTAATATTTGACTTATTTGGTGAATTTCTGCTATGAGAT 1800
 DB 1741 GGTTCATGTGTAACTTAATATTTGACTTATTTGGTGAATTTCTGCTATGAGAT 1800
 QY 1801 CTCATTTCTTTTGAATTTGCTGTGTGGATTTATTTAGAGACTGTGTCTTACTAGCA 1860
 DB 1801 CTCATTTCTTTTGAATTTGCTGTGTGGATTTATTTAGAGACTGTGTCTTACTAGCA 1860

Qy	1861	TCCTCATATATATTTTGGCTTCTTCGACATTTGGAGTCAGTTTCACTAAAGTT	1922
Db	1861	TTCTCTCATATATATTTTGGCTTCTTCGACATTTGGAGTCAGTTTCACTAAAGTT	1920
Qy	1921	TCCAATTTTATAGTAGCCTGTAAACGGAAATGTCTTACAGATTTATCTTTTATGTAT	1980
Db	1921	TCCAATTTTATAGTAGCCTGTAAACGGAAATGTCTTACAGATTTATCTTTTATGTAT	1980
Qy	1981	TTAAGTACAGAAACTAGTTTGGTGAATGTCTGTGGGTGAAGAAACAAAGTA	2040
Db	1981	TTAAGTACAGAAACTAGTTTGGTGAATGTCTGTGGGTGAAGAAACAAAGTA	2040
Qy	2041	TGCAGTAGTATATATAACATATCAAGTTGAGATCTTCTCCACATGTACTCTTTGCT	2100
Db	2041	TGCAGTAGTATATATAACATATCAAGTTGAGATCTCTCTCCACATGTACTCTTTGCT	2100
Qy	2101	CCAACTTATCATTAACACAATCCAAATTTGTGGCAATCTTTTCTCAGTGTTTTTCGGC	2160
Db	2101	CCAACTTATCATTAACACAATCCAAATTTGTGGCAATCTTTTCTCAGTGTTTTTCGGC	2160
Qy	2161	TCGAGCTTTGAGGGCAATGAGAGTTTACCTCGAATTTGGTAGACCTTGGACATCAAGT	2220
Db	2161	TCGAGCTTTGAGGGCAATGAGAGTTTACCTCGAATTTGGTAGACCTTGGACATCAAGT	2220
Qy	2221	GTTTACTAATTTGAGATTTGGAATGTATGTCTTGGGAGTTCTTGCACTCAGCTACTGCAT	2280
Db	2221	GTTTACTAATTTGAGATTTGGAATGTATGTCTTGGGAGTTCTTGCACTCAGCTACTGCAT	2280
Qy	2281	AAAAAGATATTTTCTGGATATTTCTTATGTAAATCTCACTGACAGCATATGTCTGG	2340
Db	2281	AAAAAGATATTTTCTGGATATTTCTTATGTAAATCTCACTGACAGCATATGTCTGG	2340
Qy	2341	TGGTTTCTTAAAGCCTTATTTTATAGAGAGCTTCCCTCTTGATTTGTATTTCTCT	2400
Db	2341	TGGTTTCTTAAAGCCTTATTTTATAGAGAGCTTCCCTCTTGATTTGTATTTCTCT	2400
Qy	2401	GTACTTTCTTTAACCAGATATGAATATAATGGCAAGTATCTATTTCAATGCTGGTGAACA	2460
Db	2401	GTACTTTCTTTAACCAGATATGAATATAATGGCAAGTATCTATTTCAATGCTGGTGAACA	2460
Qy	2461	CCATCTCGATGCTGCTGTATGTACAAAATCTCTGGTGGAGACTGGTATTGGAAACA	2520
Db	2461	CCATCTCGATGCTGCTGTATGTACAAAATCTCTGGTGGAGACTGGTATTGGAAACA	2520
Qy	2521	CCATGAGCATTCGACATGACCGAGGATGCTTCTGTACATATATAATTTATATATC	2580
Db	2521	CCATGAGCATTCGACATGACCGAGGATGCTTCTGTACATATATAATTTATATATC	2580
Qy	2581	ACATCTTCATATAGGCTATCAAGAAAGAAAGCAGTAGTAATAACAAAAGTTTCTATCT	2640
Db	2581	ACATCTTCATATAGGCTATCAAGAAAGAAAGCAGTAGTAATAACAAAAGTTTCTATCT	2640
Qy	2641	GTTTGGGCTATTAATTTCCCTTCCAGTTCTGTATTTATAGCTAAGGACAGACATATATTC	2700
Db	2641	GTTTGGGCTATTAATTTCCCTTCCAGTTCTGTATTTATAGCTAAGGACAGACATATATTC	2700
Qy	2701	AGATATATTTGTTACTGCTTTAACAACATGAATTAAGCTTACTTTTCAATGCTTTATAAT	2760
Db	2701	AGATATATTTGTTACTGCTTTAACAACATGAATTAAGCTTACTTTTCAATGCTTTATAAT	2760
Qy	2761	TATGATTTCTAGCTGGAGATTTCTATATATGCTGGGGTGTCTAGTGGGGTGCCTCTCTCT	2820
Db	2761	TATGATTTCTAGCTGGAGATTTCTATATATGCTGGGGTGTCTAGTGGGGTGCCTCTCTCT	2820
Qy	2821	TACACTTCTCCAGGCAATGATCTTGTGAACCAACCCCGTGAACCTCGGAATCTAGGTGCAT	2880
Db	2821	TACACTTCTCCAGGCAATGATCTTGTGAACCAACCCCGTGAACCTCGGAATCTAGGTGCAT	2880
Qy	2881	AATTTTGAATTTGTCTGAAGAAAGAAATGCTTGTGTGTATTCATTCAGATGAGAACTT	2940
Db	2881	AATTTTGAATTTGTCTGAAGAAAGAAATGCTTGTGTGTATTCATTCAGATGAGAACTT	2940

QY	2941	TTAGGATACCACTTTTATACGAAAAAAACAAAGTAAGGTTGACCTAAAGTGAAATC	3000
Db	2941	TTAGGATACCACTTTTATACGAAAAAAACAAAGTAAGGTTGACCTAAAGTGAAATC	3000
QY	3001	CCAAATTTCTAGGAATTCACCTTAACCTGTGTATTAATAATTACTGGTATCTTCATAC	3060
Db	3001	CCAAATTTCTAGGAATTCACCTTAACCTGTGTATTAATAATTACTGGTATCTTCATAC	3060
QY	3061	AACAGACGGAGAGAGTAAGGTAGATTGTGATCAAAATGCTCGTTGCATATTAATGCTC	3120
Db	3061	AACAGACGGAGAGAGTAAGGTAGATTGTGATCAAAATGCTCGTTGCATATTAATGCTC	3120
QY	3121	ATGATGTAGATTGGCACAATACGATTTCTTCTACTACTAGAAAGAAAGCTGTATGA	3180
Db	3121	ATGATGTAGATTGGCACAATACGATTTCTTCTACTACTAGAAAGAAAGCTGTATGA	3180
QY	3181	ACCACTGATGCTAGATACATCGGGACAAATATGCGTATTCGACCTAGACCTTCACTAT	3240
Db	3181	ACCACTGATGCTAGATACATCGGGACAAATATGCGTATTCGACCTAGACCTTCACTAT	3240
QY	3241	ATCCGTCCTTTGGTTGCAGTTGGCAATATACATTCCTGTTGCAGAAATTCGTGCA	3300
Db	3241	ATCCGTCCTTTGGTTGCAGTTGGCAATATACATTCCTGTTGCAGAAATTCGTGCA	3300
QY	3301	ATAAACTATGACTGCTGATATACAAAGGCAAGATTCAGAGAGACAAACGGGAAATGTTG	3360
Db	3301	ATAAACTATGACTGCTGATATACAAAGGCAAGATTCAGAGAGACAAACGGGAAATGTTG	3360
QY	3361	GTTTGGGGAAGAGCCCGCTCAAGAGTGTATCATTAATMAGCCCACTAGCTTAGTAAC	3420
Db	3361	GTTTGGGGAAGAGCCCGCTCAAGAGTGTATCATTAATMAGCCCACTAGCTTAGTAAC	3420
QY	3421	TTAGCTATATTCCTTTTCAGACCTAAACATTTTGAATGTATGAATCTCTTGCTCT	3480
Db	3421	TTAGCTATATTCCTTTTCAGACCTAAACATTTTGAATGTATGAATCTCTTGCTCT	3480
QY	3481	CGGTTTGTGGAGATGTGGGGTGTATCTACTACAAACATCGTGAAGAACTAATAATGCT	3540
Db	3481	CGGTTTGTGGAGATGTGGGGTGTATCTACTACAAACATCGTGAAGAACTAATAATGCT	3540
QY	3541	TCTCTTAACGCTGGATGGTGAAGCTATTTCTACATTTACAGTTACAAACATTAATAATG	3600
Db	3541	TCTCTTAACGCTGGATGGTGAAGCTATTTCTACATTTACAGTTACAAACATTAATAATG	3600
QY	3601	ACGTTGTTCAATCAATAGGCTTTGAATGTGACAGGTGGGATTTGGCTCGCATTTGCATT	3660
Db	3601	ACGTTGTTCAATCAATAGGCTTTGAATGTGACAGGTGGGATTTGGCTCGCATTTGCATT	3660
QY	3661	ATGTTCCCTGAATCTTAAAGTCTTTCTTCTGACCGTACCGGCTCTCTTGATTAAGTAA	3720
Db	3661	ATGTTCCCTGAATCTTAAAGTCTTTCTTCTGACCGTACCGGCTCTCTTGATTAAGTAA	3720
QY	3721	AATACTATCCATATTTGGCTATATTTGGAAGCTCTTTGCAAGTAAAGCAAGCTAAACA	3780
Db	3721	AATACTATCCATATTTGGCTATATTTGGAAGCTCTTTGCAAGTAAAGCAAGCTAAACA	3780
QY	3781	TATAAATCATCTTTGTTTAAACCAAGTCTTGGCACTACTTTCAGTCATATTTCTCA	3840
Db	3781	TATAAATCATCTTTGTTTAAACCAAGTCTTGGCACTACTTTCAGTCATATTTCTCA	3840
QY	3841	TCTTCTCTTTGATGAGCCAAAGAGACGATGACCGATCCGATCAAAGTAAGCACACAA	3900
Db	3841	TCTTCTCTTTGATGAGCCAAAGAGACGATGACCGATCCGATCAAAGTAAGCACACAA	3900
QY	3901	TCTAATATATTTATTTTACATCTCTTTGATCTTACTGTGATTTGCTGACAAATATTCG	3960
Db	3901	TCTAATATATTTATTTTACATCTCTTTGATCTTACTGTGATTTGCTGACAAATATTCG	3960
QY	3961	GGTTTGTGTTTGTAAAAAGGTATGGAATAATTTGGAAGCTATTTGTGAGAGTCAAT	4020
Db	3961	GGTTTGTGTTTGTAAAAAGGTATGGAATAATTTGGAAGCTATTTGTGAGAGTCAAT	4020
QY	4021	ACAGATCATTCGGGAATTTATGATGTACAGAGCTGTGTTCTCATTTTCTACTT	4080

```

|||||
Db 4021 ACAGGATCATCGGGGATTTATGATTGATGAGAGTGTGTTGTTCTCATTTTCTACTT 4080
Qy 4081 ATTACGTTAATGGAACCTTGAATCATCAAAAGACCGCCCAAAACCAAAATGCAAAAT 4140
Db 4081 ATTACGTTAATGGAACCTTGAATCATCAAAAGACCGCCCAAAACCAAAATGCAAAAT 4140
Qy 4141 GATGCGATAGACATTTCTTGTGATTTGTATGCTATAGCTTTCAAAATCTAGCTAGC 4200
Db 4141 GATGCGATAGACATTTCTTGTGATTTGTATGCTATAGCTTTCAAAATCTAGCTAGC 4200
Qy 4201 CTTATGATTTCCCTAGATTTATCAAAAGTTAGCCGCTTTCTAAATTTATTTATGAT 4260
Db 4201 CTTATGATTTCCCTAGATTTATCAAAAGTTAGCCGCTTTCTAAATTTATTTATGAT 4260
Qy 4261 ATTTGATTTATGCTTTCTTAATGATCAATTAAGTTGATGATCATCATCAATTTATTTTA 4320
Db 4261 ATTTGATTTATGCTTTCTTAATGATCAATTAAGTTGATGATCATCATCAATTTATTTTA 4320
Qy 4321 CTACATCAAAATGAAGTCTTTTATAGTACAGGACGACTAATGAACCTGATTAAT 4380
Db 4321 CTACATCAAAATGAAGTCTTTTATAGTACAGGACGACTAATGAACCTGATTAAT 4380
Qy 4381 ACATGATATATCAAAAGATTTATAGTGTCTTTTCAAGTTTTCATCACAAGAACTCA 4440
Db 4381 ACATGATATATCAAAAGATTTATAGTGTCTTTTCAAGTTTTCATCACAAGAACTCA 4440
Qy 4441 CCGACATTTATCAAGTTGACTTCACTTTTATGACTTAATGTTGTTGAGGATCCACAA 4500
Db 4441 CCGACATTTATCAAGTTGACTTCACTTTTATGACTTAATGTTGTTGAGGATCCACAA 4500
Qy 4501 ACTACTATTTACAGATTTTACAAAATGTGATTTCTTTATTTTATTTTCTTTGCTCAA 4560
Db 4501 ACTACTATTTACAGATTTTACAAAATGTGATTTCTTTATTTTATTTTCTTTGCTCAA 4560
Qy 4561 ACTAACATTTTCACTTCACTTTTAAAGAGAGTCCCTCCAGCTGGAAGATTTACACAAA 4620
Db 4561 ACTAACATTTTCACTTCACTTTTAAAGAGAGTCCCTCCAGCTGGAAGATTTACACAAA 4620
Qy 4621 GCTACGAGCAAAATGCTCAACAAAAGACATAGAAGGGAGCTCAGAGCTCAACCC 4680
Db 4621 GCTACGAGCAAAATGCTCAACAAAAGACATAGAAGGGAGCTCAGAGCTCAACCC 4680
Qy 4681 GAGTAGCAGCATCAAAAGTTGCTAAACAGAACCATCATTTACTTCTGAGCGGCTTTCG 4740
Db 4681 GAGTAGCAGCATCAAAAGTTGCTAAACAGAACCATCATTTACTTCTGAGCGGCTTTCG 4740
Qy 4741 TACCAAGCTCCTCAGCAAAATGTGATTTCTATGTTGATATCATGAGATCTAACTTCA 4800
Db 4741 TACCAAGCTCCTCAGCAAAATGTGATTTCTATGTTGATATCATGAGATCTAACTTCA 4800
Qy 4801 AAGTTTTTATACAAAACCTTACATTTCTTAATTTGTAGTCAAGTAAACAAATTTTCA 4860
Db 4801 AAGTTTTTATACAAAACCTTACATTTCTTAATTTGTAGTCAAGTAAACAAATTTTCA 4860
Qy 4861 TCTAATAACATTTTACTG 4879
Db 4861 TCTAATAACATTTTACTG 4879

```

RESULT 2
US-09-817-774-28

```

; Sequence 28, Application US/0981774
; Patent No. US2002012011A1
; GENERAL INFORMATION:
; APPLICANT: CHOE, Sung-hwa
; APPLICANT: FELDMANN A., Kenneth
; TITLE OF INVENTION: DWTs MUTANTS
; FILE REFERENCE: 2225-0020 / 91020.002
; CURRENT APPLICATION NUMBER: US/09/817,774
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/192,202
; PRIOR FILING DATE: 2000-03-27

```

```

; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 28
; LENGTH: 4880
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (671)..(757)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (1050)..(1271)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (1340)..(1396)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (1505)..(1606)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (1711)..(1776)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (2174)..(2287)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (2418)..(2546)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (2776)..(2874)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (3259)..(3384)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (3494)..(3559)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (3636)..(3716)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (3806)..(3889)
; OTHER INFORMATION: dwf5
; NAME/KEY: CDS
; LOCATION: (3982)..(4047)
; OTHER INFORMATION: dwf5
US-09-817-774-28

Query Match 99.8%; Score 4868; DB 10; Length 4880;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4879; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 AACTGTCCTTCCTCCCAACAGCATCTCCCTTTCGGGCTGCTAGCATTTTCGTTAAG 60
Db 1 AACTGTCCTTCCTCCCAACAGCATCTCCCTTTCGGGCTGCTAGCATTTTCGTTAAG 60
Qy 61 ATGTTGATTCGTCGAATTAATCTGTTCCCTTCTATCATCATCATCATCATCATCAT 120
Db 61 ATGTTGATTCGTCGAATTAATCTGTTCCCTTCTATCATCATCATCATCATCATCAT 120
Qy 121 ATTCGATCTCTTAACATCTAATAAATTTCTGGAATGAAAAAATCTGAGAAACGCAAA 180
Db 121 ATTCGATCTCTTAACATCTAATAAATTTCTGGAATGAAAAAATCTGAGAAACGCAAA 180
Qy 181 GAATTCACACAGATCAGAAATCTAGTCTAGTTTCGGCGGAGTTACCAAAACATTT 240
Db 181 GAATTCACACAGATCAGAAATCTAGTCTAGTTTCGGCGGAGTTACCAAAACATTT 240
Qy 241 AGAAAAATCATCTTACAGATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAG 300
Db 241 AGAAAAATCATCTTACAGATTCAGCAATTCAGCAATTCAGCAATTCAGCAATTCAG 300
Qy 301 CCATTGCTTGGAGTAGAAGAAATGTGTAATGATGATGATGATGATGATGATGATGAT 360

```

Db 301 CCAATGCTCGAGTAAGAAGAAATGTAACTATGACGCGCAATGAGATGAAGAA 360
Qy 361 TTTGATTTTGTGTCGATCATATGTTGTAATACAGATACAGAAATCTT 420
Db 361 TTTGATTTTGTGTCGATCATATGTTGTAATACAGATACAGAAATCTT 420
Qy 421 TTGAGGTTGAGAGAGAGAGAGAGAGATGGATCAATCAAAAGATGGCAGAA 480
Db 421 TTGAGGTTGAGAGAGAGAGAGAGATGGATCAATCAAAAGATGGCAGAA 480
Qy 481 AAAGTCTCAAGGTTAAATTTGAATTAAGAAAAACATGTGAGAGACGTGACAGTGG 540
Db 481 AAAGTCTCAAGGTTAAATTTGAATTAAGAAAAACATGTGAGAGACGTGACAGTGG 540
Qy 541 ATTATCTGAGTTTCACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 600
Db 541 ATTATCTGAGTTTCACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 600
Qy 601 ACTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Db 601 ACTCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy 661 GAAGAGAAATATGCGAGAGAGTACATTCGAGATGTTACTTACAGATGATGATGAT 720
Db 661 GAAGAGAAATATGCGAGAGAGTACATTCGAGATGTTACTTACAGATGATGATGAT 720
Qy 721 GCTTCGCGCTTCTGTCACAGTCTTCTGTCATCTCTGTCATCTCTGTCATCTCT 780
Db 721 GCTTCGCGCTTCTGTCACAGTCTTCTGTCATCTCTGTCATCTCTGTCATCTCT 780
Qy 781 CAAGTCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Db 781 CAAGTCTTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
Qy 841 GCGAGAGTGCATAGCGCATATCTTAAATCTTCCATTAAGTGAAGTGAAGAG 900
Db 841 GCGAGAGTGCATAGCGCATATCTTAAATCTTCCATTAAGTGAAGTGAAGAG 900
Qy 901 TCTTTATTTGAAAAAGCATATCATATCTTCTGCGCAATGTGATTTCCAGATTT 960
Db 901 TCTTTATTTGAAAAAGCATATCATATCTTCTGCGCAATGTGATTTCCAGATTT 960
Qy 961 TTGAGAAACAAGGATGATTAAGCTACAAATCATATTTAGTGGTTTAAATAGC 1020
Db 961 TTGAGAAACAAGGATGATTAAGCTACAAATCATATTTAGTGGTTTAAATAGC 1020
Qy 1021 TCACAGTTTCATTTCTTTATGATGAGATGATGATGATGATGATGATGATGAT 1080
Db 1021 TCACAGTTTCATTTCTTTATGATGAGATGATGATGATGATGATGATGATGAT 1080
Qy 1081 TTTACTAGACCTTTGGCTTTTGGGAGATGAGATGATGATGATGATGATGATGAT 1140
Db 1081 TTTACTAGACCTTTGGCTTTTGGGAGATGAGATGATGATGATGATGATGATGAT 1140
Qy 1141 CAAGACCCACTTGTATGCTTGAATTAATTTTGTATGAGATGATGATGATGATGAT 1200
Db 1141 CAAGACCCACTTGTATGCTTGAATTAATTTTGTATGAGATGATGATGATGATGAT 1200
Qy 1201 TTTAGCTGCTTGTGCTGTTAAAGAGTGAAGGTCATATCTCCAGCGGAAACCGAC 1260
Db 1201 TTTAGCTGCTTGTGCTGTTAAAGAGTGAAGGTCATATCTCCAGCGGAAACCGAC 1260
Qy 1261 CAGTTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1261 CAGTTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Qy 1321 TAGGTTCTTTCTTGTGAGGCAATGATGATGATGATGATGATGATGATGATGAT 1380
Db 1321 TAGGTTCTTTCTTGTGAGGCAATGATGATGATGATGATGATGATGATGATGAT 1380
Qy 1381 TCTTGGCTTTTGTGATGAAGAAATGTTCCGATGATGATGATGATGATGATGATGAT 1440
Db 1381 TCTTGGCTTTTGTGATGAAGAAATGTTCCGATGATGATGATGATGATGATGATGAT 1440

Db 1381 TCTTGGCTTTTGTGATGAAGAAATGTTCCGATGATGATGATGATGATGATGATGAT 1440
Qy 1441 TTAGATCTTAAAGTGTGATTAAGTGTGACACCTATGATGATTAATTTCTGCGTAT 1500
Db 1441 TTAGATCTTAAAGTGTGATTAAGTGTGACACCTATGATGATTAATTTCTGCGTAT 1500
Qy 1501 CAGTTTGAATCTTCAACCCGCAATGATGATGATGATGATGATGATGATGATGAT 1560
Db 1501 CAGTTTGAATCTTCAACCCGCAATGATGATGATGATGATGATGATGATGATGAT 1560
Qy 1561 ACTAATATTCGAGACCTTCAATTTTGTGTTGTTGATGATGATGATGATGATGATGAT 1620
Db 1561 ACTAATATTCGAGACCTTCAATTTTGTGTTGTTGATGATGATGATGATGATGATGAT 1620
Qy 1621 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Db 1621 CATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
Qy 1681 TTTGCTAAACCTATTTTCTTCTGATGAGGCGCATGTTGACCTTCAATGATGATCT 1740
Db 1681 TTTGCTAAACCTATTTTCTTCTGATGAGGCGCATGTTGACCTTCAATGATGATCT 1740
Qy 1741 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Db 1741 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
Qy 1801 CTCATTTCTTTTGAATTTCTGTTGGATTTATTTGAGAGATGATGATGATGATGAT 1860
Db 1801 CTCATTTCTTTTGAATTTCTGTTGGATTTATTTGAGAGATGATGATGATGATGAT 1860
Qy 1861 TCTCTCAATATTTTGGCTTCTTCTGACATTTGAGAGTCAAGTTTACTAAAGTT 1920
Db 1861 TCTCTCAATATTTTGGCTTCTTCTGACATTTGAGAGTCAAGTTTACTAAAGTT 1920
Qy 1921 TCCAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
Db 1921 TCCAAATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
Qy 1981 TTAAGTACAGAAATCTGTTTGGTGAATGCTGCTGGTGAAGAGAAAGACAAAGTA 2040
Db 1981 TTAAGTACAGAAATCTGTTTGGTGAATGCTGCTGGTGAAGAGAAAGACAAAGTA 2040
Qy 2041 TGCAGTATGATTAATTAACAAATGATGATGATGATGATGATGATGATGATGAT 2100
Db 2041 TGCAGTATGATTAATTAACAAATGATGATGATGATGATGATGATGATGATGAT 2100
Qy 2101 CCAACTTATCATTAACAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Db 2101 CCAACTTATCATTAACAAATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
Qy 2161 TGTGCTTTGACAGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db 2161 TGTGCTTTGACAGGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Qy 2221 GTTACTAATTCAGATTCGGAATGATGATGATGATGATGATGATGATGATGAT 2280
Db 2221 GTTACTAATTCAGATTCGGAATGATGATGATGATGATGATGATGATGATGAT 2280
Qy 2281 AAAAGAGTATTTCTGATTAATTTCTTATGTAATTAATTAATTAATTAATTAATTAAT 2340
Db 2281 AAAAGAGTATTTCTGATTAATTTCTTATGTAATTAATTAATTAATTAATTAAT 2340
Qy 2341 TGTGTTCTTTAAGTCTTATTTCAATTAAGAGCTTCACTCTTCTGATTTGATTTCT 2400
Db 2341 TGTGTTCTTTAAGTCTTATTTCAATTAAGAGCTTCACTCTTCTGATTTGATTTCT 2400
Qy 2401 GTACTTCTTTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
Db 2401 GTACTTCTTTTACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2460
Qy 2461 CCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520
Db 2461 CCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2520

[illegible]

OY	3601	ACGTGTCACATCAATGSCCTTGAATGCAAGGTGGGGAATGGCTGCAATTCATT	3660
Db	3601	ACGTTGTCATCAATCAATGGCTTGAATGCAAGGTGGGGAATGGCTGCAATTCATT	3660
OY	3661	ATGTTCCGTGAGACTTAAAGTGGCTTCTTGCGACGTACCGGCTCTCTGCATAGCTAA	3720
Db	3661	ATGTTCCGTGAGACTTAAAGTGGCTTCTTGCGACGTACCGGCTCTCTGCATAGCTAA	3720
OY	3721	AATACATCCATCTATTGGTATATTCGAAGCTCTTCTGCAGTGAAGCCACTTAACA	3780
Db	3721	AATACATCCATCTATTGGTATATTCGAAGCTCTTCTGCAGTGAAGCCACTTAACA	3780
OY	3781	TATTAATCACTTGTGTTAAACAGAGTCTTGGCATACCTCTACGCATATTTTCACGCT	3840
Db	3781	TATTAATCACTTGTGTTAAACAGAGTCTTGGCATACCTCTACGCATATTTTCACGCT	3840
OY	3841	TCTTCTCTTTGATCGAGCCAAAGAGACATGACGATGCCGATCAAAAGTAAAGCAACAA	3900
Db	3841	TCTTCTCTTTGATCGAGCCAAAGAGACATGACGATGCCGATCAAAAGTAAAGCAACAA	3900
OY	3901	TCTAATGATTTATTTTACACATCCCTTTGATCTTACTCTGATTTGCTGACAATATTCG	3960
Db	3901	TCTAATGATTTATTTTACACATCCCTTTGATCTTACTCTGATTTGCTGACAATATTCG	3960
OY	3961	GGTTTTGTTTTGGTAAAAAGCTATGGGAATATATGGAACCTGATTTGAG- AAGTCAAA	4019
Db	3961	GGTTTTGTTTTGGTAAAAAGCTATGGGAATATATGGAACCTGATTTGAG- AAGTCAAA	4020
OY	4020	TACAGATCATCCCGGGAATTTATGATTTGAACGAAGCTGTGTCATTTCTACT	4079
Db	4021	TACAGATCATTCGGGAATTTATGATTTGAACGAAGCTGTGTCATTTCTACT	4080
OY	4080	TATTACGTTAATTCGAACGTTGGAAATCATCAAAAAGCGTGCCAAAACAAAAATGCAAT	4139
Db	4081	TATTACGTTAATTCGAACGTTGGAAATCATCAAAAAGCGTGCCAAAACAAAAATGCAAT	4140
OY	4140	TGATGCGATGACACTTCTTTGCTGATTTGATAGCATAGTGGTTTCCAATCCTAGCTAC	4199
Db	4141	TGAATCGATGACACTTCTTTGCTGATTTGATAGCATAGTGGTTTCCAATCCTAGCTAC	4200
OY	4200	GCTATGATATTCCTCGATTAATCAAGTTAGCGTCCGTTTCTCAATTTATATTGGA	4259
Db	4201	GCTATGATATTCCTCGATTAATCAAGTTAGCGTCCGTTTCTCAATTTATATTGGA	4260
OY	4260	TATTTGATTATACGCTCTTATATGATCATCAATAAGTTGATCATTCATTAATTTT	4319
Db	4261	TATTTGATTATACGCTCTTATATGATCATCAATAAGTTGATCATTCATTAATTTT	4320
OY	4320	ACTATCATCAATGAAAGCTTTTAAAGTACAGGCACTATTTCAACTGAAACCCATATAAA	4379
Db	4321	ACTATCATCAATGAAAGCTTTTAAAGTACAGGCACTATTTCAACTGAAACCCATATAAA	4380
OY	4380	TACATGATATATCAAAAGAAATGATAGTGTCTTCAGTTTCTCAATCACAAAAGACTC	4439
Db	4381	TACATGATATATCAAAAGAAATGATAGTGTCTTCAGTTTCTCAATCACAAAAGACTC	4440
OY	4440	ACCGACATTAATCAAGTGACTCACTTTAGACTTAATGATGTGCTCGAGAGATCCAAA	4499
Db	4441	ACCGACATTAATCAAGTGACTCACTTTAGACTTAATGATGTGCTCGAGAGATCCAAA	4500
OY	4500	AACACTATATACAGAGTTTCCACAAAATGTGATCTTTATTTTATTTTCTTTGGTCA	4559
Db	4501	AACACTATATACAGAGTTTCCACAAAATGTGATCTTTATTTTATTTTCTTTGGTCA	4560
OY	4560	AACATCAATTCATCAAACTTTAAAGAGAGTGGCTCCACGTGGAGAGTATACCAAAA	4619
Db	4561	AACATCAATTCATCAAACTTTAAAGAGAGTGGCTCCACGTGGAGAGTATACCAAAA	4620
OY	4620	AGCTACAGAGCAAAATCTCAAAACAAAAGACATAAGAAAGGGAGCTCAGAACTCAAC	4679
Db	4621	AGCTACAGAGCAAAATCTCAAAACAAAAGACATAAGAAAGGGAGCTCAGAACTCAAC	4680
OY	4680	CGAGTAGACGATCCAAAGTTGCTAAAAACAAACCATCATTTACTTCTGAGCGGCTTTC	4739


```

1144 CAAGACCCACTTGGATTGCTGGAATAATATATTTGCTATGAGACATTTGAAGCTATTC 1200
Qy 1201 TTCAGTGTCTGCTGCTGGTAAAGAGTGGAGGTCCAAATATCTCCAGCGGAAACGAC 1260
Db 1201 TTCAGTGTCTGCTGCTGGTAAAGAGTGGAGGTCCAAATATCTCCAGCGGAAACGAC 1260
Qy 1261 CAGTTTACAGAGTATGTATAGTCTGTGTGGAGAGTGTGAGAGTGAAGACAGT 1320
Db 1261 CAGTTTACAGAGTATGTATAGTCTGTGTGGAGAGTGTGAGAGTGAAGACAGT 1320
Qy 1321 TACGTTCTTTCTTCTGACGCGCAAGTGTGCGCTTACTTTGTGACACTAGCAACCTA 1380
Db 1321 TACGTTCTTTCTTCTGACGCGCAAGTGTGCGCTTACTTTGTGACACTAGCAACCTA 1380
Qy 1381 TCTTGTCTTTGGTGTAGAAATGTTTCCAGAGTCTGTTGGTTGTTAACTTATGTT 1440
Db 1381 TCTTGTCTTTGGTGTAGAAATGTTTCCAGAGTCTGTTGGTTGTTAACTTATGTT 1440
Qy 1441 TACGATCTTAAAGCTTGATTAAGTGTACACCTATGTAGTTAAATTTCTGCTTAT 1500
Db 1441 TACGATCTTAAAGCTTGATTAAGTGTACACCTATGTAGTTAAATTTCTGCTTAT 1500
Qy 1501 CAGGTTTGAATCTTCAACCTGCAATGTCTATGATCACTTGGTGAATATTTTCGCG 1560
Db 1501 CAGGTTTGAATCTTCAACCTGCAATGTCTATGATCACTTGGTGAATATTTTCGCG 1560
Qy 1561 ACTAATATTCGGAAGCTTATATTTGTGTTTGTGTGACATTAAGTGAAGTCTATG 1620
Db 1561 ACTAATATTCGGAAGCTTATATTTGTGTTTGTGTGACATTAAGTGAAGTCTATG 1620
Qy 1621 CATGAATGAGTGTCTTCTTGTCTTCTATCTATTTTGTAGAAATTTGACCTTTGTC 1680
Db 1621 CATGAATGAGTGTCTTCTTGTCTTCTATCTATTTTGTAGAAATTTGACCTTTGTC 1680
Qy 1681 TTGCTAAAACCTATTTTTTCTTGTGACAGGCGCATGTTGACACTCAAGTACTCT 1740
Db 1681 TTGCTAAAACCTATTTTTTCTTGTGACAGGCGCATGTTGACACTCAAGTACTCT 1740
Qy 1741 GGTTCATGAGTGTACCTAATATGACTCTATGTTGGTGAAGTATTTTGTGATGACAT 1800
Db 1741 GGTTCATGAGTGTACCTAATATGACTCTATGTTGGTGAAGTATTTTGTGATGACAT 1800
Qy 1801 CTCATCTTTTGTGAATGCTGTTGGATTAATTTGAGAGAACTGCTGCTTACTAGCCA 1860
Db 1801 CTCATCTTTTGTGAATGCTGTTGGATTAATTTGAGAGAACTGCTGCTTACTAGCCA 1860
Qy 1861 TCTCTCATATATTTTTTGGCTCTTCTGACATTTGAGAGTCAAGTTTCACTAAAGTT 1920
Db 1861 TCTCTCATATATTTTTTGGCTCTTCTGACATTTGAGAGTCAAGTTTCACTAAAGTT 1920
Qy 1921 TCGAATTTTAGTAGACCTGTAACGGAATGCTTGAAGAGTATTAATCTTTTATGAT 1980
Db 1921 TCGAATTTTAGTAGACCTGTAACGGAATGCTTGAAGAGTATTAATCTTTTATGAT 1980
Qy 1981 TTAAGTACAGAAACCTAGTTTGTGATGTCTGCTGAGAGGAAAGAAAGAAAGTA 2040
Db 1981 TTAAGTACAGAAACCTAGTTTGTGATGTCTGCTGAGAGGAAAGAAAGAAAGTA 2040
Qy 2041 TGCAGTAGTATTAATCAATATCAAGTTGAGATCTTCTCAACCATGTACCTTTTGTCT 2100
Db 2041 TGCAGTAGTATTAATCAATATCAAGTTGAGATCTTCTCAACCATGTACCTTTTGTCT 2100
Qy 2101 CCAACTATCATTAACAAATCAATTTGGGCAATCTTTTCACTGCTGTTTTCGCG 2160
Db 2101 CCAACTATCATTAACAAATCAATTTGGGCAATCTTTTCACTGCTGTTTTCGCG 2160
Qy 2161 TCGTGTCTTGCAGGAGTGTATGACCTGCAATTTGTAAGAGCTTTGACATCAAGT 2220
Db 2161 TCGTGTCTTGCAGGAGTGTATGACCTGCAATTTGTAAGAGCTTTGACATCAAGT 2220
Qy 2221 GTTACTAATTCAGATTCGGAATGATGCTTGGGCAATCTTTCGACGTCAGTACTGAT 2280
Db 2221 GTTACTAATTCAGATTCGGAATGATGCTTGGGCAATCTTTCGACGTCAGTACTGAT 2280

```

```

Db 2221 GTTACTAATTCAGATTCGGAATGATGCTTGGGCAATCTTTCGACGTCAGTACTGAT 2280
Qy 2281 AAAACAGGATTTTCTGAGTAATTTCTATGTAATATCTACACTGTGACATAGTCTGG 2340
Db 2281 AAAACAGGATTTTCTGAGTAATTTCTATGTAATATCTACACTGTGACATAGTCTGG 2340
Qy 2341 TGGTTTCTTTAAGTCTTATTTTCAATAGAGTCTTCACTCTTCTGATTTGATTTCT 2400
Db 2341 TGGTTTCTTTAAGTCTTATTTTCAATAGAGTCTTCACTCTTCTGATTTGATTTCT 2400
Qy 2401 GTACTTTCTTTAACCAGTATGAATTAATGGAAGATCTGATTCATCTGCTGAGCA 2460
Db 2401 GTACTTTCTTTAACCAGTATGAATTAATGGAAGATCTGATTCATCTGCTGAGCA 2460
Qy 2461 CCATCCGATGCTGCTGATGTACAAATCTTCTGCTGGAAGCTGTTATGAGACA 2520
Db 2461 CCATCCGATGCTGCTGATGTACAAATCTTCTGCTGGAAGCTGTTATGAGACA 2520
Qy 2521 CCATGACATTTGCACATGACGAGTATGCTTCTGTACATATATAAATTTATCATC 2580
Db 2521 CCATGACATTTGCACATGACGAGTATGCTTCTGTACATATATAAATTTATCATC 2580
Qy 2581 ACATCTCATATAGGATACAAAGAAAGAAAGCCAGTATATACAAAGTCTCTACT 2640
Db 2581 ACATCTCATATAGGATACAAAGAAAGAAAGCCAGTATATACAAAGTCTCTACT 2640
Qy 2641 GTTGGGCTATTAATTTCTGTCAGTCTGTTATATAGGTAAGGACAGCATATTATTC 2700
Db 2641 GTTGGGCTATTAATTTCTGTCAGTCTGTTATATAGGTAAGGACAGCATATTATTC 2700
Qy 2701 AGATTAATTTGCTACGCTTTAACACATGAATTAAGCTTACTTTCAATGCTTTATAT 2760
Db 2701 AGATTAATTTGCTACGCTTTAACACATGAATTAAGCTTACTTTCAATGCTTTATAT 2760
Qy 2761 TATGATTTCTTACGCTGATCTATATATGCTGGGGTGTCTAGTGTGGGCTCTGTC 2820
Db 2761 TATGATTTCTTACGCTGATCTATATATGCTGGGGTGTCTAGTGTGGGCTCTGTC 2820
Qy 2821 TACACTTCTCCAGCATGTACTTGTGAACACCCCTGCAACTCGGAACTCAGAGTGCAT 2880
Db 2821 TACACTTCTCCAGCATGTACTTGTGAACACCCCTGCAACTCGGAACTCAGAGTGCAT 2880
Qy 2881 AATTTGAAATTTGCTAGAGAAAGAAATGCTTTGTGTTAATCAATGAGAACTT 2940
Db 2881 AATTTGAAATTTGCTAGAGAAAGAAATGCTTTGTGTTAATCAATGAGAACTT 2940
Qy 2941 TTAGATATCCACTTTTATTCGAAAAAAACAGTAAAGTGTAGCTTAATGGAATC 3000
Db 2941 TTAGATATCCACTTTTATTCGAAAAAAACAGTAAAGTGTAGCTTAATGGAATC 3000
Qy 3001 CCAATTTCTAGAAATGCACTTCTGTTATTTGAATTAAGTGTATCTTCCATTACA 3060
Db 3001 CCAATTTCTAGAAATGCACTTCTGTTATTTGAATTAAGTGTATCTTCCATTACA 3060
Qy 3061 AACAGACGGAGAGATGAAGTATGATTTGTATCAAAATGCTGTTGATATATGCTC 3120
Db 3061 AACAGACGGAGAGATGAAGTATGATTTGTATCAAAATGCTGTTGATATATGCTC 3120
Qy 3121 ATGATTTAGTGTGCAAGTATCACTTTTCTAGTAGTAAGAAAGGCTTCAATGAAAC 3180
Db 3121 ATGATTTAGTGTGCAAGTATCACTTTTCTAGTAGTAAGAAAGGCTTCAATGAAAC 3180
Qy 3181 ACCACTGATGTACATGATGAGGACATATGCTATCTCGTACCTAAGTCTCACTAT 3240
Db 3181 ACCACTGATGTACATGATGAGGACATATGCTATCTCGTACCTAAGTCTCACTAT 3240
Qy 3241 ATCCGCTTTTGGTTCAGTTGGCAATATATCTTCTGTCAGGAATCTGTGCAATTTAC 3300
Db 3241 ATCCGCTTTTGGTTCAGTTGGCAATATATCTTCTGTCAGGAATCTGTGCAATTTAC 3300
Qy 3301 ATAACTATGACTGTGATAGCAAAAGCAAGAGTTCAGAGGACAAACGGGAAATGTTG 3360
Db 3301 ATAACTATGACTGTGATAGCAAAAGCAAGAGTTCAGAGGACAAACGGGAAATGTTG 3360

```

QY	3361	GTTTGGGGAAGAGCCCGCTCAAGAGTGTGATCATATATAGCCCACTAGTTAGTGAAC	3420
Db	3361	TTTTGGGGGAAGAGCCCGCTCAAGAGTGTGATCATATATAGCCCACTAGTTAGTGAAC	3420
QY	3421	TTAGCATATTCCTTTTCAGACCTTAACATATTTGATGTGTATGTAATCTCTTGSTGCT	3480
Db	3421	TTAGCATATATTCCTTTTCAGACCTTAACATATTTGATGTGTATGTAATCTCTTGSTGCT	3480
QY	3481	CGGTTTGTGCGAGATTTGGCGCTGTATATCAACAACATCTGGTGAACCTAAACTAGTCT	3540
Db	3481	CGGTTTGTGCGAGATTTGGCGCTGTATATCAACAACATCTGGTGAACCTAAACTAGTCT	3540
QY	3541	TCTCTTAACGCTGGATGGTGTGAGTACTATTCTACANTATACAGTTTACAACCTATAAATAATAG	3600
Db	3541	TCTCTTAACGCTGGATGGTGTGAGTACTATTCTACANTATACAGTTTACAACCTATAAATAATAG	3600
QY	3601	ACGTTGTCATCANTCAATAGGCTTTGTAATGTGACAGTGGGAGTTGGCTCGTATTTCCATT	3660
Db	3601	ACGTTGTCATCANTCAATAGGCTTTGTAATGTGACAGTGGGAGTTGGCTCGTATTTCCATT	3660
QY	3661	ATGTTCCTGACATCTTAAAGTGGCTTTCTTGTGACCGTACCGGCTCTCTTGCATTAAGCTAA	3720
Db	3661	ATGTTCCTGACATCTTAAAGTGGCTTTCTTGTGACCGTACCGGCTCTCTTGCATTAAGCTAA	3720
QY	3721	AATFATTCACATCTATTTGGTATATTTGAGACCTCTCTCTTGAAGTGAAGAACCACTTACA	3780
Db	3721	AATFATTCACATCTATTTGGTATATTTGAGACCTCTCTCTTGAAGTGAAGAACCACTTACA	3780
QY	3781	TATTAATCATCTTTGTTTAAACCAAGTCTTGTGACATCTTCTACGTCATATTTCTCACCCCT	3840
Db	3781	TATTAATCATCTTTGTTTAAACCAAGTCTTGTGACATCTTCTACGTCATATTTCTCACCCCT	3840
QY	3841	TCTTCTCTTGAAGAGCAAGAGAGAGATGACCGATGCGCATCAAAAGTAAAGCACCAAA	3900
Db	3841	TCTTCTCTTGAAGAGCAAGAGAGAGATGACCGATGCGCATCAAAAGTAAAGCACCAAA	3900
QY	3901	TCTAATGATTTATTTTACACATCCTCTTGTATCTACTGCTGATTTGCTGCATATTTTCG	3960
Db	3901	TCTAATGATTTATTTTACACATCCTCTTGTATCTACTGCTGATTTGCTGCATATTTTCG	3960
QY	3961	GGTTTGTGTTTGGTAAAAAGGTATGGGAAATATTTGGAAGCTGTATGTGTGAG-AAAGTCAAA	4019
Db	3961	GGTTTGTGTTTGGTAAAAAGGTATGGGAAATATTTGGAAGCTGTATGTGTGAGAAAGTCAAA	4020
QY	4020	TACAGATCATTCGCGGAATTTATTTGATTGTAGAGAGTGTGTGTCTCTATTTTCTACT	4079
Db	4021	TACAGATCATTCGCGGAATTTATTTGATTGTAGAGAGTGTGTGTCTCTATTTTCTACT	4080
QY	4080	TATTAAGTTAATTCGAAAGCTTGAATCATCAAAAGACCGTGCCCAAAACAAAATGCAAAAT	4139
Db	4081	TATTAAGTTAATTCGAAAGCTTGAATCATCAAAAGACCGTGCCCAAAACAAAATGCAAAAT	4140
QY	4140	TGATGCGGATACACATCTTTTGTGATTTGTATGCTATAGTGTTCAAATCTCTAGCTAC	4199
Db	4141	TGATGCGGATACACATCTTTTGTGATTTGTATGCTATAGTGTTCAAATCTCTAGCTAC	4200
QY	4200	GCTTATGATTTCCCTAGATTTATCAAAAGTTAGCTGCGGCTTTCTAATTTTATTTATTTGA	4259
Db	4201	GCTTATGATTTCCCTAGATTTATCAAAAGTTAGCTGCGGCTTTCTAATTTTATTTATTTGA	4260
QY	4260	TATTTTGAATTTACTGCTTCTTAATGCAATCAAAATAGTGTGATTCATCATCTAATATTTT	4319
Db	4261	TATTTTGAATTTACTGCTTCTTAATGCAATCAAAATAGTGTGATTCATCATCTAATATTTT	4320
QY	4320	ACTACATCAAAATGAGTCTTTTATAGTACAGGACATAATTTGAACCTGAAGCCATATAAA	4379
Db	4321	ACTACATCAAAATGAGTCTTTTATAGTACAGGACATAATTTGAACCTGAAGCCATATAAA	4380
QY	4380	TACATGAATTAATCACAAAAGATGTATAGTGTCTTTTCAGTTTTCATCAACAAGAACTC	4439
Db	4381	TACATGAATTAATCACAAAAGATGTATAGTGTCTTTTCAGTTTTCATCAACAAGAACTC	4440

QY	4440	ACCGACATTAATACAGTTTGAAGTTCACATTTTGAACCTTAATTTGATGCTGTGAGATGTCACAA	4499
Db	4441	ACCGACATTAATACAGTTTGAAGTTCACATTTTGAACCTTAATTTGATGCTGTGAGATGTCACAA	4500
QY	4500	AATCTATTATACAGATTTTCACAAAATGTGATTCCTTATTTTATTTTCTTTGGTCA	4559
Db	4501	AATCTATTATTAACAGATTTTCACAAAATGTGATTCCTTATTTTATTTTCTTTGGTCA	4560
QY	4560	AATTAACATTTTCAATTCATCACTTTTAAACGAGAGTTGCTCCACGTGAGAGTATACACAAA	4619
Db	4561	AATTAACATTTTCAATTCATCACTTTTAAACGAGAGTTGCTCCACGTGAGAGTATACACAAA	4620
QY	4620	AGCTACGAGACAAATGCTCAACAAAAGACATTAAGAAAGGAGCTCAGAACTCAACC	4679
Db	4621	AGCTACGAGACAAATGCTCAACAAAAGACATTAAGAAAGGAGAGCTCAGAAAGCTCAACC	4680
QY	4680	CGAGTAGCAGCATCCCAAAATGCTTAAACAGAACCATCATTTACTTCTGAGGCGGTCTTC	4739
Db	4681	CGAGTAGCAGCATCCCAAAATGCTTAAACAGAACCATCATTTACTTCTGAGGCGGTCTTC	4740
QY	4740	GTACCAAGCTCCTTCAGCAAAATTTGATTTCTATGTGTTGGATATCATGATCTAATTC	4799
Db	4741	GTACCAAGCTCCTTCAGCAAAATTTGATTTCTATGTGTTGGATATCATGATCTAATTC	4800
QY	4800	AAAGTTTATTAATACAAAACCTTACATTTCTTAATTTGTAGTCAAACTAACAATTTTC	4859
Db	4801	AAAGTTTATTAATACAAAACCTTACATTTCTTAATTTGTAGTCAAACTAACAATTTTC	4860
QY	4860	ATCTAAAAACAATTTTACTG 4879	
Db	4861	ATCTAAAAACAATTTTACTG 4880	

RESULT 4

US-09-817-774-34

; Sequence 34, Application US/09817774

; Patent No. US2002012011A1

; GENERAL INFORMATION:

; APPLICANT: CHOE, Sunghwa

; TITLE OF INVENTION: Dwf5 MUTANTS

; FILE REFERENCE: 2225-0020 / 91020.002

; CURRENT APPLICATION NUMBER: US/09/817,774

; PRIORITY FILING DATE: 2001-03-26

; PRIOR APPLICATION NUMBER: 60/1192,202

; NUMBER OF SEQ ID NOS: 45

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 34

; LENGTH: 4880

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (671)..(757)

; OTHER INFORMATION: dwf5-2

; NAME/KEY: CDS

; LOCATION: (1050)..(1271)

; OTHER INFORMATION: dwf5-2

; NAME/KEY: CDS

; LOCATION: (1340)..(1396)

; OTHER INFORMATION: dwf5-2

; NAME/KEY: CDS

; LOCATION: (1505)..(1606)

; OTHER INFORMATION: dwf5-2

; NAME/KEY: CDS

; LOCATION: (1711)..(1776)

; OTHER INFORMATION: dwf5-2

; NAME/KEY: CDS

; LOCATION: (2174)..(2287)

; OTHER INFORMATION: dwf5-2

; NAME/KEY: CDS

; LOCATION: (2418)..(2546)

OTHER INFORMATION: dwf5-2
 NAME/KEY: CDS
 LOCATION: (2776)..(2874)
 OTHER INFORMATION: dwf5-2
 NAME/KEY: CDS
 LOCATION: (3321)..(3386)
 OTHER INFORMATION: dwf5-2
 NAME/KEY: CDS
 LOCATION: (3496)..(3525)
 OTHER INFORMATION: dwf5-2
 US-09-817-774-34

Query Match 99.7%; Score 4866.4; DB 10; Length 4880;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4878; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AACTGTCCTTCCTCCACACACAGATCTCTTCCTCCGGCTGCTACGATTTGTTAAGG 60
 DB 1 AACTGTCCTTCCTCCACACACAGATCTCTTCCTCCGGCTGCTACGATTTGTTAAGG 60
 QY 61 ATGTGATTCGTCGATAAATCTGTCCTTCTATCATCATCATCAGTACAGGAGCTA 120
 DB 61 ATGTGATTCGTCGATAAATCTGTCCTTCTATCATCATCATCAGTACAGGAGCTA 120
 QY 121 ATTCGATCTCTAAACATCTCTAAATTTCTGGAATGAAAAAACTGAGAAACGCAAA 180
 DB 121 ATTCGATCTCTAAACATCTCTAAATTTCTGGAATGAAAAAACTGAGAAACGCAAA 180
 QY 181 GAATTCACACACAGCATCAGAAAACTAGATCTAGGTTTGGCGGAGTTTCCAAAAACAT 240
 DB 181 GAATTCACACACAGCATCAGAAAACTAGATCTAGGTTTGGCGGAGTTTCCAAAAACAT 240
 QY 241 AGAAAAATCGACTACTTACAGCAATTCAGCAATTCGAGAAATTTCTCCAGGCAAG 300
 DB 241 AGAAAAATCGACTACTTACAGCAATTCAGCAATTCGAGAAATTTCTCCAGGCAAG 300
 QY 301 CCATTGCTTCGAGTAAGAAGAAATGTAACATATGACTGCGAATCAGAGTGAAGAGA 360
 DB 301 CCATTGCTTCGAGTAAGAAGAAATGTAACATATGACTGCGAATCAGAGTGAAGAGA 360
 QY 361 TTTGATTTTGTAGTCTCATTCATTCACATGTTGTAATACAGAGTACCAATCTT 420
 DB 361 TTTGATTTTGTAGTCTCATTCATTCACATGTTGTAATACAGAGTACCAATCTT 420
 QY 421 TTGAGGTTGAGAGAGAGAGTAAGATTTGGATATCAACAAAGAGTGGGAAGA 480
 DB 421 TTGAGGTTGAGAGAGAGAGTAAGATTTGGATATCAACAAAGAGTGGGAAGA 480
 QY 481 AAAGTCTCAAGGGTAAATTTGAAATTAAGAAAAACATGTGAGAGCAGCTGACGTGG 540
 DB 481 AAAGTCTCAAGGGTAAATTTGAAATTAAGAAAAACATGTGAGAGCAGCTGACGTGG 540
 QY 541 ATTATCTAGATTTCACAGTAGACCTGTGAAGTCAATGCTCTTCATCGTTAACTTCG 600
 DB 541 ATTATCTAGATTTCACAGTAGACCTGTGAAGTCAATGCTCTTCATCGTTAACTTCG 600
 QY 601 ACTCTGTTGTTAGTCAACAGATCAGATCTGAGGCTTTGGCGAGAGAGAGAAGA 660
 DB 601 ACTCTGTTGTTAGTCAACAGATCAGATCTGAGGCTTTGGCGAGAGAGAGAAGA 660
 QY 661 GAAGAAGAAAAAGCGAGAGACTGTACATTCGATGCTTACTTACGCAATGATTTATC 720
 DB 661 GAAGAAGAAAAAGCGAGAGACTGTACATTCGATGCTTACTTACGCAATGATTTATC 720
 QY 721 GCTTCGCTTCCTGTCACACTTTCCTCATCTCTGTAAGCTCAATCAATTTCTAATTC 780
 DB 721 GCTTCGCTTCCTGTCACACTTTCCTCATCTCTGTAAGCTCAATCAATTTCTAATTC 780
 QY 781 CAAGTCTTATTTAGTCTCAGATCAGACTGCGACACTTTCTGCTCCTTAATTCATA 840
 DB 781 CAAGTCTTATTTAGTCTCAGATCAGACTGCGACACTTTCTGCTCCTTAATTCATA 840
 QY 841 GCGAGAAGTGACATAGCCGATATCTTAAATCTTTCCCATTAAGTAGACTTGATTAAG 900

DB 841 GCGAGAAGTGACATAGCCGATATCTTAAATCTTTCCCATTAAGTAGACTTGATTAAG 900
 QY 901 TCTTTAATTAAGAAAGGATTCATATGATTTCTCTGCGCAATGTGAGTTCCAGATT 960
 DB 901 TCTTTAATTAAGAAAGGATTCATATGATTTCTCTGCGCAATGTGAGTTCCAGATT 960
 QY 961 TTAGAAGCAAGGATAGATTAAAGCTACAAATCATATTTAGTGTGTTAAGTAATGC 1020
 DB 961 TTAGAAGCAAGGATAGATTAAAGCTACAAATCATATTTAGTGTGTTAAGTAATGC 1020
 QY 1021 TCACAGTTTCATTTCTTTATGATGAGATGTACACAAATGTTTCAATCAGAGTGTCTG 1080
 DB 1021 TCACAGTTTCATTTCTTTATGATGAGATGTACACAAATGTTTCAATCAGAGTGTCTG 1080
 QY 1081 TTACTGAGACCTTTGGCTCTCTTTGGAGAGATGAGATTCAGAGACTATTCATATTCG 1140
 DB 1081 TTACTGAGACCTTTGGCTCTCTTTGGAGAGATGAGATTCAGAGACTATTCATATTCG 1140
 QY 1141 CAAGACCACCTTTGATGCTTGGAAAAATTTATTTGCTATGAGAGCATTTGAAGTATTC 1200
 DB 1141 CAAGACCACCTTTGATGCTTGGAAAAATTTATTTGCTATGAGAGCATTTGAAGTATTC 1200
 QY 1201 TTCAGTCTCTTCGCTGCTGTAAGAAAGAGTTGAGGTCATATCTCCAGCCGAAACGAC 1260
 DB 1201 TTCAGTCTCTTCGCTGCTGTAAGAAAGAGTTGAGGTCATATCTCCAGCCGAAACGAC 1260
 QY 1261 CAGTTTCAAGATATCTTCAATTAAGTCTGTGTGAGAGAGTGAATGAGATGACAGAC 1320
 DB 1261 CAGTTTCAAGATATCTTCAATTAAGTCTGTGTGAGAGAGTGAATGAGATGACAGAC 1320
 QY 1321 TACGCTCTCTTTCTGAGGCAATGCTGCTGCTTACTTTGTGACATACCAACCTA 1380
 DB 1321 TACGCTCTCTTTCTGAGGCAATGCTGCTGCTTACTTTGTGACATACCAACCTA 1380
 QY 1381 TCTTGCTTTTGGTGGTAAGAAATGTTCCGATGACTTGGTTGGTTAAGTATTTGT 1440
 DB 1381 TCTTGCTTTTGGTGGTAAGAAATGTTCCGATGACTTGGTTGGTTAAGTATTTGT 1440
 QY 1441 TACGATCTTAAAGCTTGATTAAGTGTGACACCTATGATGATTAATTTCTGCTGAT 1500
 DB 1441 TACGATCTTAAAGCTTGATTAAGTGTGACACCTATGATGATTAATTTCTGCTGAT 1500
 QY 1501 CAGGTTGGAATCTTCAACCCCTGCATTTGCTATGATCACTTGGGTGAATTTTCGGC 1560
 DB 1501 CAGGTTGGAATCTTCAACCCCTGCATTTGCTATGATCACTTGGGTGAATTTTCGGC 1560
 QY 1561 ACTAATATTCGGAAGCTTCATATTTTGTGTTTGTGTATATAAAGTAAAGTGTATTC 1620
 DB 1561 ACTAATATTCGGAAGCTTCATATTTTGTGTTTGTGTATATAAAGTAAAGTGTATTC 1620
 QY 1621 CATGATGCTGCTTCTTGTGTTTCTTAGTCTCAATTTTGTGAATTTGACCTTTGTC 1680
 DB 1621 CATGATGCTGCTTCTTGTGTTTCTTAGTCTCAATTTTGTGAATTTGACCTTTGTC 1680
 QY 1681 TTGCTAAAACTTATTTTCTTGTACAGGCGCATTTGACCTTCATCAAGTACTCT 1740
 DB 1681 TTGCTAAAACTTATTTTCTTGTACAGGCGCATTTGACCTTCATCAAGTACTCT 1740
 QY 1741 GGTTCATGCTGTAACCTAATATTTGATCTGATTTGGGTGAGTATTTCTGCTATCAGAT 1800
 DB 1741 GGTTCATGCTGTAACCTAATATTTGATCTGATTTGGGTGAGTATTTCTGCTATCAGAT 1800
 QY 1801 CTCATTTCTTTTGAATTTGCTTGTGGGATTTATTTAGAGAACTGCTGTTACTAGCCA 1860
 DB 1801 CTCATTTCTTTTGAATTTGCTTGTGGGATTTATTTAGAGAACTGCTGTTACTAGCCA 1860
 QY 1861 TCTCCTCATTAATTTTGTCTCTTCTGCACTTTGAGAGTCAAGTTTCACTAAAGTT 1920
 DB 1861 TCTCCTCATTAATTTTGTCTCTTCTGCACTTTGAGAGTCAAGTTTCACTAAAGTT 1920
 QY 1921 TCCAAATTTAGTACGCTGTAAAGGAAATGTTTGTAGCAATTTATCTTTTATTTGAT 1980

Db 1921 TCCAAATTTAGTAGACCTGTAACGGAAGATGCTTACAGATTTATCTTTTATGAT 1980
Qy 1981 TTAAGTACAGAAAGACTAGTTGGTGTGAATGCTGTGCGGGAAGAAAGCAAAAGTA 2040
Db 1991 TTAAGTACAGAAAGACTAGTTGGTGTGAATGCTGTGCGGGAAGAAAGCAAAAGTA 2040
Qy 2041 TGCAGTAGTATTTAACAATATCAATGATGATCTTCTGCACAGTACTCTTTTGGT 2100
Db 2041 TGCAGTAGTATTTAACAATATCAATGATGATCTTCTGCACAGTACTCTTTTGGT 2100
Qy 2101 CCACATTAATCATTAACCAATCCAAATTTGGGCAATCTTTTCTCAGTGTGTTTTCGG 2160
Db 2101 CCACATTAATCATTAACCAATCCAAATTTGGGCAATCTTTTCTCAGTGTGTTTTCGG 2160
Qy 2161 TCGTGTCTTTCAGGAGGAGTGTGACCTCGAATGGTAAGGCTTTACATCAAGT 2220
Db 2161 TCGTGTCTTTCAGGAGGAGTGTGACCTCGAATGGTAAGGCTTTACATCAAGT 2220
Qy 2221 GTTTACTAATTTGAGATTCGGAATGATGCTTGGGCAAGTCTTGCACTGACATGATGAT 2280
Db 2221 GTTTACTAATTTGAGATTCGGAATGATGCTTGGGCAAGTCTTGCACTGACATGATGAT 2280
Qy 2281 AAAACAGATATTTCTGATATATTTCTATGTAATTAATCTACACTGTGACATAGTCTGG 2340
Db 2281 AAAACAGATATTTCTGATATATTTCTATGTAATTAATCTACACTGTGACATAGTCTGG 2340
Qy 2341 TGGTTCCTTAAGTCTTATTTCTATGAGAGGCTTCACCTCTTTCGATTTGATTCCT 2400
Db 2341 TGGTTCCTTAAGTCTTATTTCTATGAGAGGCTTCACCTCTTTCGATTTGATTCCT 2400
Qy 2401 GTACTTTCTTTAACCAATGAAATTAATGCAAAAGTATGATCAATCTGCTGATGACA 2460
Db 2401 GTACTTTCTTTAACCAATGAAATTAATGCAAAAGTATGATCAATCTGCTGATGACA 2460
Qy 2461 CCATCCGATGCTGCTGTATGTCACAAATTTCTGCTGGAAGCTGTTATTTGACA 2520
Db 2461 CCATCCGATGCTGCTGTATGTCACAAATTTCTGCTGGAAGCTGTTATTTGACA 2520
Qy 2521 CCATGACATTTGACATGACGAGATGCTTCTGATACATATTAATTAATTTATCATC 2580
Db 2521 CCATGACATTTGACATGACGAGATGCTTCTGATACATATTAATTAATTTATCATC 2580
Qy 2581 ACATCTCATATAGGATACAGAAAGAAAGCAAGTATTAACAAAGTTTCTCTACT 2640
Db 2581 ACATCTCATATAGGATACAGAAAGAAAGCAAGTATTAACAAAGTTTCTCTACT 2640
Qy 2641 GTTTGGCTATATTTCTCTGCCAGTCTCTTTATATGCTTAAGGACGACATATATTC 2700
Db 2641 GTTTGGCTATATTTCTCTGCCAGTCTCTTTATATGCTTAAGGACGACATATATTC 2700
Qy 2701 AGATTAATTTAGTGGCTTTAACAATGAACTTAAGCTTCTTCAATCTTTATTAAT 2760
Db 2701 AGATTAATTTAGTGGCTTTAACAATGAACTTAAGCTTCTTCAATCTTTATTAAT 2760
Qy 2761 TATGATTTCTTCACTGATTTCTATATATGCTGGGTTGTCTAGTGGGGTCCCTTCGTC 2820
Db 2761 TATGATTTCTTCACTGATTTCTATATATGCTGGGTTGTCTAGTGGGGTCCCTTCGTC 2820
Qy 2821 TACACTCTGACGAGATGATGCTTGTGAACACCCGTCGAATCGGAATCTAGGTGAT 2880
Db 2821 TACACTCTGACGAGATGATGCTTGTGAACACCCGTCGAATCGGAATCTAGGTGAT 2880
Qy 2881 AATTTGAATTTGTCTAGAGAAAGAAATGCTTTGTGTGTTAATCCATGGAATCT 2940
Db 2881 AATTTGAATTTGTCTAGAGAAAGAAATGCTTTGTGTGTTAATCCATGGAATCT 2940
Qy 2941 TTAGGATACCACTTTTATACGAAAAAACAAGTAAAGTGAAGTAAATGGAATCT 3000
Db 2941 TTAGGATACCACTTTTATACGAAAAAACAAGTAAAGTGAAGTAAATGGAATCT 3000
Qy 3001 CCATTTCTAGGAATGCACTTAACCTGTTATTAATTAATCTGATCTTCCATTTACA 3060
Db 3001 CCATTTCTAGGAATGCACTTAACCTGTTATTAATTAATCTGATCTTCCATTTACA 3060

Qy 3061 AACAGCGGAGAGAGATTAAGTATGATTTGATCAAAATGCTCGTGCATATTTAGCTC 3120
Db 3061 AACAGCGGAGAGAGATTAAGTATGATTTGATCAAAATGCTCGTGCATATTTAGCTC 3120
Qy 3121 ATGATTTAGTGGCAGATATGATGTTTCTAGTGTATACAAAGGCTTATGATTAAC 3180
Db 3121 ATGATTTAGTGGCAGATATGATGTTTCTAGTGTATACAAAGGCTTATGATTAAC 3180
Qy 3181 ACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Db 3181 ACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Qy 3241 ATCCGCTTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Db 3241 ATCCGCTTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Qy 3301 ATAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
Db 3301 ATAACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3360
Qy 3361 GTTTGGGGAAGAGCCCGTCAAAAGTGTGATCATATAAGCCCACTAGCTTATGAC 3420
Db 3361 GTTTGGGGAAGAGCCCGTCAAAAGTGTGATCATATAAGCCCACTAGCTTATGAC 3420
Qy 3421 TTAGCTATATTTCTTTGACACTTAAACATTTGATGATGATGATGATGATGATGAT 3480
Db 3421 TTAGCTATATTTCTTTGACACTTAAACATTTGATGATGATGATGATGATGATGAT 3480
Qy 3481 CGGTTTGTGACAGATTTGCGCTGCTGATACATCAATCTGATGATGATGATGATGAT 3540
Db 3481 CGGTTTGTGACAGATTTGCGCTGCTGATACATCAATCTGATGATGATGATGATGAT 3540
Qy 3541 TCTCTTACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
Db 3541 TCTCTTACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
Qy 3601 ACCTGTTATATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
Db 3601 ACCTGTTATATCATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3660
Qy 3661 ATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
Db 3661 ATGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3720
Qy 3721 AATTAATCATATTTGCTATATTTGCAAGCTCTTTCGCAAGTGAAGCCACTTAACA 3780
Db 3721 AATTAATCATATTTGCTATATTTGCAAGCTCTTTCGCAAGTGAAGCCACTTAACA 3780
Qy 3781 TATTAATCATATTTGCTATATTTGCAAGCTCTTTCGCAAGTGAAGCCACTTAACA 3840
Db 3781 TATTAATCATATTTGCTATATTTGCAAGCTCTTTCGCAAGTGAAGCCACTTAACA 3840
Qy 3841 TCTTCTCTTGTATGAGCCAGAGAGAGATGATGATGATGATGATGATGATGATGAT 3900
Db 3841 TCTTCTCTTGTATGAGCCAGAGAGAGATGATGATGATGATGATGATGATGATGAT 3900
Qy 3901 TCTTCTCTTGTATGAGCCAGAGAGAGATGATGATGATGATGATGATGATGATGAT 3960
Db 3901 TCTTCTCTTGTATGAGCCAGAGAGAGATGATGATGATGATGATGATGATGATGAT 3960
Qy 3961 GGTGTTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4020
Db 3961 GGTGTTGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4020
Qy 4020 TACAGATCATTCGGAATTTATGATTTATGATTTATGATTTATGATTTATGATTT 4080
Db 4020 TACAGATCATTCGGAATTTATGATTTATGATTTATGATTTATGATTTATGATTT 4080
Qy 4080 TATTAAGTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4139
Db 4081 TATTAAGTATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4140

QY	4140	TTGGCAGATGACATTCCTTTTGGTGAATTTGATAGCATAGTGGTTTTCAAATCTGATGAC	4139
Db	4141	TGATGCGATGACATCTTTTGGCTGATTTGTATGCTATAGGTTTTCAAATCTGATGAC	4200
QY	4200	GCTTATGTATTTCCCTAGATTATCAAAAGTTAGCCGCCGTTTCTAAATTTATTTTGA	4259
Db	4201	GCTTATGATTTCCCTAGATTATCAAAAGTTAGCCGCCGTTTCTAAATTTATTTTGA	4260
QY	4260	TATTTTGAATTAAGTCTGTTCTTAATGATCAAAATAAGTTGATTCAATCTAATATTTT	4319
Db	4261	TATTTTGAATTAAGTCTGTTCTTAATGATCAAAATAAGTTGATTCAATCTAATATTTT	4320
QY	4320	ACTGATCACAATGAGAGCTTTTAAAGTACAGGACCTAAATFAGACTTGAACCCCTAATAA	4379
Db	4321	ACTGATCACAATGAGAGCTTTTAAAGTACAGGACCTAAATFAGACTTGAACCCCTAATAA	4380
QY	4380	TACATGATATATCAAAAGAATGATAGTGTCTTCAAGTTTTCATACAAAGATC	4439
Db	4381	TACATGATATATCAAAAGAATGATAGTGTCTTCAAGTTTTCATACAAAGATC	4440
QY	4440	ACCGACATTTATCAAGTTGACTTCACTTTTAGACTTAATTTGGTGTCTGAGATCCACA	4499
Db	4441	ACCGACATTTATCAAGTTGACTTCACTTTTAGACTTAATTTGGTGTCTGAGATCCACA	4500
QY	4500	AACGACATTAATCAAGTTTTCACAAAATGATGATCTTATTTTATTTTCTTTGGTCA	4559
Db	4501	AACGACATTAATCAAGTTTTCACAAAATGATGATCTTATTTTATTTTCTTTGGTCA	4560
QY	4560	AACGACATTTTCATCAACTTTAAAGGAGAGTTGCCGCCACGATGAGATATACAAA	4619
Db	4561	AACGACATTTTCATCAACTTTAAAGGAGAGTTGCCGCCACGATGAGATATACAAA	4620
QY	4620	AGCTACGAGCAATGCTCAAAACAAAAGACATAAGAAAGGAGAGCTCAAGAGCTCAAC	4679
Db	4621	AGCTACGAGCAATGCTCAAAACAAAAGACATAAGAAAGGAGAGCTCAAGAGCTCAAC	4680
QY	4680	CGAGTAGAGATCCAAAGTTGCTAAAGACGACACATTAATTAATCTCTGAGGCGGTCC	4739
Db	4681	CGAGTAGAGATCCAAAGTTGCTAAAGACGACACATTAATTAATCTCTGAGGCGGTCC	4740
QY	4740	GTACCAAGCTCCTTCAGCAAAATTTGATTTCTATGTTTGATATCATGATCTAATCTC	4799
Db	4741	GTACCAAGCTCCTTCAGCAAAATTTGATTTCTATGTTTGATATCATGATCTAATCTC	4800
QY	4800	AAAGTTTTTAATTAACAAAACCTTACATTTCTTAATTTGTAGTCAAAAGCTAACAATTC	4859
Db	4801	AAAGTTTTTAATTAACAAAACCTTACATTTCTTAATTTGTAGTCAAAAGCTAACAATTC	4860
QY	4860	ATCTAAAAACAATTTTACTG 4879	
Db	4861	ATCTAAAAACAATTTTACTG 4880	

```

1  RESULT 5
2  US-09-817-774-36
3  : Sequence 36, Application US/09817774
4  : Patent No. US2002012011A1
5  :
6  : GENERAL INFORMATION:
7  :
8  : APPLICANT: CHOIE, Sung-hwa
9  :
10 : APPLICANT: FELDMANN A., Kenneth
11 :
12 : TITLE OF INVENTION: DWT5 MUTANTS
13 :
14 : FILE REFERENCE: 2225-0020 / 91020. 002
15 :
16 : CURRENT APPLICATION NUMBER: US/09/817, 774
17 :
18 : CURRENT FILING DATE: 2001-03-26
19 :
20 : PRIOR APPLICATION NUMBER: 60/192,202
21 :
22 : PRIOR FILING DATE: 2000-03-27
23 :
24 : NUMBER OF SEQ ID NOS: 45
25 :
26 : SOFTWARE: PatentIn Ver. 2.0
27 :
28 : SEQ ID NO 36
29 :
30 : LENGTH: 4880
31 :
32 : TYPE: DNA
33 :
34 : ORGANISM: Arabidopsis thaliana
35 :
36 : FEATURE:

```

```

NAME/KEY: CDS
LOCATION: (671)..(757)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (1050)..(1271)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (1340)..(1396)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (1505)..(1606)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (1711)..(1776)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (2174)..(2287)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (2418)..(2546)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (2776)..(2874)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (3259)..(3384)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (3494)..(3559)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (3636)..(3716)
OTHER INFORMATION: dwf5-3
NAME/KEY: CDS
LOCATION: (3806)..(3856)
OTHER INFORMATION: dwf5-3
US-09-817-774-36

```

	Query Match	Similarity	99.7%;	Score	4866.4;	DB	10;	Length	4880;
	Best Local	Similarity	100.0%;	Pred.	No. 0;				
	Matches	4878;	Conservative	0;	Mismatches	1;	Indels	1;	Gaps
OY	1	AACGTTCCTTCCTCCACACACAGATCTCTCTTCCGGCTGCTACGATTTGCTTAAAG	60						
Db	1	AACGTTCCTTCCTCCACACACAGATCTCTCTTCCGGCTGCTACGATTTGCTTAAAG	60						
OY	61	ATGTGATTCGCGAATAAATCTGTCTCTTCTATCATCATCATCATAGGACACTA	120						
Db	61	ATGTGATTCGCGAATAAATCTGTCTCTTCTATCATCATCATCATAGGACACTA	120						
OY	121	ATTGCATCTCTTAAACATCTAATAATTTCTGGAATGAAAAAAACCTGGAAGCCAAA	180						
Db	121	ATTGCATCTCTTAAACATCTAATAATTTCTGGAATGAAAAAAACCTGGAAGCCAAA	180						
OY	181	GAATTCACACACGATCCAGAAAACAGATCTGTGGTTCCGCGGAGTTACCAAAAACATT	240						
Db	181	GAATTCACACACGATCCAGAAAACAGATCTGTGGTTCCGCGGAGTTACCAAAAACATT	240						
OY	241	AGAAAAAATCGATCTACAGCAATTCGGAATTTCTTCCAAAGGCAG	300						
Db	241	AGAAAAAATCGATCTACAGCAATTCGGAATTTCTTCCAAAGGCAG	300						
OY	301	CCATTCGTTCCGAGTAAGAAATGGTAACCTATGATCGGATTCAGAGTGAAGAGA	360						
Db	301	CCATTCGTTCCGAGTAAGAAATGGTAACCTATGATCGGATTCAGAGTGAAGAGA	360						
OY	361	TTTGGATTTTGTAGTCTCGATTCACTATGTGTTGAATACAGAGATGACGAAGTCTT	420						
Db	361	TTTGGATTTTGTAGTCTCGATTCACTATGTGTTGAATACAGAGATGACGAAGTCTT	420						
OY	421	TTGAGGGTTGAGAGAGAAAGCTTAAGAAATGGGATTAATCAAAAAGATGGGGAATA	480						
Db	421	TTGAGGGTTGAGAGAGAAAGCTTAAGAAATGGGATTAATCAAAAAGATGGGGAATA	480						

QY 481 AAAGTCTCAAGGGTAAATTTGGAATTAGAAAAACATGTGGAGAGACAGTGAACGTGGG 540
DB 481 AAAGTCTCAAGGGTAAATTTGGAATTAGAAAAACATGTGGAGAGACAGTGAACGTGGG 540
QY 541 ATTTATCTGAGATTTCACAGTAGACCTGTGAAGTTCATGTCCTTCACTGTTAACTTCGAG 600
DB 541 ATTTATCTGAGATTTCACAGTAGACCTGTGAAGTTCATGTCCTTCACTGTTAACTTCGAG 600
QY 601 ACTGCTGATTTAGTCAACACAGATCAATCTGAGGTTGGCCGAGAGAGAGAGCA 660
DB 601 ACTGCTGATTTAGTCAACACAGATCAATCTGAGGTTGGCCGAGAGAGAGAGCA 660
QY 661 GAAGAGAAATGGGAGAGACTAGATTCGCGATCGTACTTACGATGAGATGATTC 720
DB 661 GAAGAGAAATGGGAGAGACTAGATTCGCGATCGTACTTACGATGAGATGATTC 720
QY 721 GCTTCGCTCTTCTGTCACCTTTGTCATTCCTCTGTAAGCTCATGATTTTGATTCG 780
DB 721 GCTTCGCTCTTCTGTCACCTTTGTCATTCCTCTGTAAGCTCATGATTTTGATTCG 780
QY 781 CAAGCTTATTTAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 840
DB 781 CAAGCTTATTTAGTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTC 840
QY 841 GCGAGAAAGTCCATGACGATTCCTTTAAATCCTTCCATTAAGTAGAGATTCATAGAG 900
DB 841 GCGAGAAAGTCCATGACGATTCCTTTAAATCCTTCCATTAAGTAGAGATTCATAGAG 900
QY 901 TCTTTAATTTGAAAAAGCGATTCATGATTCCTTCTGCGCAATGTAGTTTCCAGATTT 960
DB 901 TCTTTAATTTGAAAAAGCGATTCATGATTCCTTCTGCGCAATGTAGTTTCCAGATTT 960
QY 961 TTAGAGAAACAGGGTAGATTTAAAGCTACAAATCATATTTAGTAGATTTAGTAGATTC 1020
DB 961 TTAGAGAAACAGGGTAGATTTAAAGCTACAAATCATATTTAGTAGATTTAGTAGATTC 1020
QY 1021 TCACAGTTTCATTTCTTTATGATGAGATGTCACATGATGTCATGAGATGTCATG 1080
DB 1021 TCACAGTTTCATTTCTTTATGATGAGATGTCACATGATGTCATGAGATGTCATG 1080
QY 1081 TTACTCAGACCTTTGCTCTTTTGGAGAGATGAGATTCAGAGATTCATGATGATG 1140
DB 1081 TTACTCAGACCTTTGCTCTTTTGGAGAGATGAGATTCAGAGATTCATGATGATG 1140
QY 1141 CAAGACCCACTTTGATTTGCTGGAATTTATTTTGTCTATGAGAGATTTGAGATTC 1200
DB 1141 CAAGACCCACTTTGATTTGCTGGAATTTATTTTGTCTATGAGAGATTTGAGATTC 1200
QY 1201 TTCAGCTGCTTCTGCTGTAAGAGATGAGATTCAGATTCATGATGATGATG 1260
DB 1201 TTCAGCTGCTTCTGCTGTAAGAGATGAGATTCAGATTCATGATGATGATG 1260
QY 1261 CAGTTTCAAGGATGATGATTTAGTCTTGTGTGTGAGAGATGATGATGATGATG 1320
DB 1261 CAGTTTCAAGGATGATGATTTAGTCTTGTGTGTGAGAGATGATGATGATGATG 1320
QY 1321 TACGCTCTTTCTTCTGAGGCAATGCTGCTGCTTACTTGTGACATGATGATGATG 1380
DB 1321 TACGCTCTTTCTTCTGAGGCAATGCTGCTGCTTACTTGTGACATGATGATGATG 1380
QY 1381 TCTTGTGCTTGTGTGTAAGAAATGTTTCCGATGATGATGATGATGATGATGATG 1440
DB 1381 TCTTGTGCTTGTGTGTAAGAAATGTTTCCGATGATGATGATGATGATGATGATG 1440
QY 1441 TACGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
DB 1441 TACGATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
QY 1501 CAGTTTGAATCTTCAACCTGCAATGCTATGATGATGATGATGATGATGATGATG 1560
DB 1501 CAGTTTGAATCTTCAACCTGCAATGCTATGATGATGATGATGATGATGATGATG 1560

QY 1561 ACTAATTTGGAGAGCTTCATATTTTGTGTTGTGATCATTAAGATGATGATG 1620
DB 1561 ACTAATTTGGAGAGCTTCATATTTTGTGTTGTGATCATTAAGATGATGATG 1620
QY 1621 CATGATGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1680
DB 1621 CATGATGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1680
QY 1681 TTGCTAAACTTATTTTTTCTTGTACAGAGGCAATGTCACCTTCATCAAGTACT 1740
DB 1681 TTGCTAAACTTATTTTTTCTTGTACAGAGGCAATGTCACCTTCATCAAGTACT 1740
QY 1741 GGTTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
DB 1741 GGTTCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1800
QY 1801 CTGATCTTTTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
DB 1801 CTGATCTTTTGAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1860
QY 1861 TCTCTCATTAATTTTTTGTGCTTTCTGACATTTGAGAGTCAAGTTTCACTAAAGT 1920
DB 1861 TCTCTCATTAATTTTTTGTGCTTTCTGACATTTGAGAGTCAAGTTTCACTAAAGT 1920
QY 1921 TCCATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980
DB 1921 TCCATTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1980
QY 1981 TTAGTACAAAGAACTAGTTGCTGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
DB 1981 TTAGTACAAAGAACTAGTTGCTGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040
QY 2041 TGCAGTAGTATTAATTAACATATCAAGTGTGAGATCTTCTGACATGATCTTTTGT 2100
DB 2041 TGCAGTAGTATTAATTAACATATCAAGTGTGAGATCTTCTGACATGATCTTTTGT 2100
QY 2101 CCACCTTATCATTAACACATCCAAATTTGTGCAATCTTCTGACATGATCTTTTGT 2160
DB 2101 CCACCTTATCATTAACACATCCAAATTTGTGCAATCTTCTGACATGATCTTTTGT 2160
QY 2161 TCGTCTTTGACAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
DB 2161 TCGTCTTTGACAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATG 2220
QY 2221 GTTACTAATTTGCAATTTGGAATGATGATGATGATGATGATGATGATGATGATG 2280
DB 2221 GTTACTAATTTGCAATTTGGAATGATGATGATGATGATGATGATGATGATGATG 2280
QY 2281 AAACAGGATTTTCTGATTAATTTCTTATGTAATTAATTAATTAATTAATTAATTA 2340
DB 2281 AAACAGGATTTTCTGATTAATTTCTTATGTAATTAATTAATTAATTAATTAATTA 2340
QY 2341 TGGTTTCTTTAAGTCTTATTTCTATTTAGAGACTTCACTCTTCTGATTTGATGATG 2400
DB 2341 TGGTTTCTTTAAGTCTTATTTCTATTTAGAGACTTCACTCTTCTGATTTGATGATG 2400
QY 2401 GTACTTCTTTAACAAGATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2460
DB 2401 GTACTTCTTTAACAAGATGTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2460
QY 2461 CCATCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
DB 2461 CCATCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
QY 2521 CCATGACATTTGACATGACGAGGATGCTCTGCTGATATTAATTAATTAATTAATTA 2580
DB 2521 CCATGACATTTGACATGACGAGGATGCTCTGCTGATATTAATTAATTAATTAATTA 2580
QY 2581 ACATCTCTATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
DB 2581 ACATCTCTATTAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
QY 2641 GTTGGCTATTAATTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATG 2700

|||||
Db 2641 GTTGGCTATATATTCCTGGCAGTTCGTTATATGCTAGAGGAGAGATATATTC 2700
Oy 2701 AGATTATTTGTAAGCTTTTAAACATGTAAGTAACTTAACTTTTCAATGCTTTATAT 2760
Db 2701 AGATTATTTGTAAGCTTTTAAACATGTAAGTAACTTAACTTTTCAATGCTTTATAT 2760
Oy 2761 TATGATTTCTGAGTGTGATATATATGCTGGGCTGTAGTGTGGTGGCTTGTGTC 2820
Db 2761 TATGATTTCTGAGTGTGATATATATGCTGGGCTGTAGTGTGGTGGCTTGTGTC 2820
Oy 2821 TACACTTCCAGCAGTACCTTGTGAACACACCCGTCGAACCTGGAACCTAGGTGAT 2880
Db 2821 TACACTTCCAGCAGTACCTTGTGAACACACCCGTCGAACCTGGAACCTAGGTGAT 2880
Oy 2881 AATTTTGAATTTGCTAGAGAAAGAAATCGTCTTGTGTGTTAATCCAATGAGAACTT 2940
Db 2881 AATTTTGAATTTGCTAGAGAAAGAAATCGTCTTGTGTGTTAATCCAATGAGAACTT 2940
Oy 2941 TTAGGATACCACTTTTATATAGAAAGAAAGAAAGTAAAGTGAAGTGAAGTGAAGT 3000
Db 2941 TTAGGATACCACTTTTATATAGAAAGAAAGAAAGTAAAGTGAAGTGAAGTGAAGT 3000
Oy 3001 CCAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3060
Db 3001 CCAATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3060
Oy 3061 AACAGACGGGAGAGAGATTAAGTGTATGTTGATCAAAATGCTGTGATATATATATG 3120
Db 3061 AACAGACGGGAGAGATTAAGTGTATGTTGATCAAAATGCTGTGATATATATATG 3120
Oy 3121 ATGATGTAGTGTGAGAT 3180
Db 3121 ATGATGTAGTGTGAGAT 3180
Oy 3181 ACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Db 3181 ACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3240
Oy 3241 ATCCGCTTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Db 3241 ATCCGCTTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Oy 3301 ATAAAT 3360
Db 3301 ATAAAT 3360
Oy 3361 GTTGGGAG 3420
Db 3361 GTTGGGAG 3420
Oy 3421 TTAGCTAT 3480
Db 3421 TTAGCTAT 3480
Oy 3481 CGGTTTGTGACAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
Db 3481 CGGTTTGTGACAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3540
Oy 3541 TCTCTTAACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
Db 3541 TCTCTTAACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
Oy 3601 AGCTTGTAT 3660
Db 3601 AGCTTGTAT 3660
Oy 3661 ATGTTCCTGATCTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3720
Db 3661 ATGTTCCTGATCTTAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 3720
Oy 3721 AAT 3780
Db 3721 AAT 3780

|||||
Db 3721 AAT 3780
Oy 3781 TAT 3840
Db 3781 TAT 3840
Oy 3841 TCTCTTCTGATGAG 3900
Db 3841 TCTCTTCTGATGAG 3900
Oy 3901 TCTAT 3960
Db 3901 TCTAT 3960
Oy 3961 GCTTTGTTTGGTAAAG 4020
Db 3961 GCTTTGTTTGGTAAAG 4020
Oy 4020 TACAGAT 4079
Db 4020 TACAGAT 4079
Oy 4080 TAT 4139
Db 4080 TAT 4139
Oy 4140 TAT 4199
Db 4140 TAT 4199
Oy 4141 TAT 4259
Db 4141 TAT 4259
Oy 4200 GCTTAT 4260
Db 4200 GCTTAT 4260
Oy 4260 TAT 4319
Db 4260 TAT 4319
Oy 4320 ACTAT 4379
Db 4320 ACTAT 4379
Oy 4380 TAT 4439
Db 4380 TAT 4439
Oy 4440 ACCGAT 4499
Db 4440 ACCGAT 4499
Oy 4500 AACTAT 4559
Db 4500 AACTAT 4559
Oy 4560 AACTAT 4619
Db 4560 AACTAT 4619
Oy 4620 AGCTAT 4679
Db 4620 AGCTAT 4679
Oy 4680 CGAGTAT 4739
Db 4680 CGAGTAT 4739
Oy 4740 GTACAT 4799
Db 4740 GTACAT 4799
Oy 4800 AAGTTTAT 4859
Db 4800 AAGTTTAT 4859
Oy 4801 AAGTTTAT 4860
Db 4801 AAGTTTAT 4860

QY 4860 ATCTAAACAACTTTACTG 4879
 Db 4861 ATCTAAACAACTTTACTG 4880

RESULT 6
 US-09-817-774-38
 : Sequence 38, Application US/09817774
 : Patent No. US2002012011A1
 : GENERAL INFORMATION:
 : APPLICANT: CHOE, Sunghwa
 : APPLICANT: FELDMANN A., Kenneth
 : TITLE OF INVENTION: DWFS MUTANTS
 : FILE REFERENCE: 2225-0020 / 91020_002
 : CURRENT APPLICATION NUMBER: US/09/817,774
 : PRIOR FILING DATE: 2001-03-26
 : PRIOR APPLICATION NUMBER: 60/192,202
 : NUMBER OF SEQ ID NOS: 45
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 38
 : LENGTH: 4880
 : TYPE: DNA
 : ORGANISM: Arabidopsis thaliana
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (671)..(757)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (1050)..(1271)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (11340)..(1396)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (1505)..(1606)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (1711)..(1776)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (2174)..(2287)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (2418)..(2546)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (2776)..(2874)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (3259)..(3384)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (3494)..(3559)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (3636)..(3716)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (3806)..(3889)
 : OTHER INFORMATION: dwfs-4
 : NAME/KEY: CDS
 : LOCATION: (3982)..(4047)
 : OTHER INFORMATION: dwfs-4
 : US-09-817-774-38

Query Match 99.7%; Score 4866.4; DB 10; Length 4880;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4878; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AACTGTCCTTCCTCCACACAGATGCTCTTTCCGGCTGCTAGATTTGCTTAAG 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db	1	AACTGTCCTTCCTCCACACAGATGCTCTTTCCGGCTGCTAGATTTGCTTAAG	60
QY	61	ATGTTATTCGTCGAATAAATCTGCTTTCTATCATCATCATCATGAGCAGTA	120
Db	61	ATGTTATTCGTCGAATAAATCTGCTTTCTATCATCATCATCATGAGCAGTA	120
QY	121	ATTGATCTCTTAAACATCTAAATTTCTGGAATGAAAAAACTGGAAGCCAAA	180
Db	121	ATTGATCTCTTAAACATCTAAATTTCTGGAATGAAAAAACTGGAAGCCAAA	180
QY	181	GAATTCACACAGCATCAGAAAACTAGATCTAGGTTTCGCGGAGTTACCAAAACATT	240
Db	181	GAATTCACACAGCATCAGAAAACTAGATCTAGGTTTCGCGGAGTTACCAAAACATT	240
QY	241	AGAAAAATCGATCTTACACGAATTCACCAATTCGTGAGAAATTTCTCCAGGCAG	300
Db	241	AGAAAAATCGATCTTACACGAATTCACCAATTCGTGAGAAATTTCTCCAGGCAG	300
QY	301	CCATTGCTTGGCGATGAAGAAATTTGTAACCTATGACTGGCAATCGAGTGAAGAA	360
Db	301	CCATTGCTTGGCGATGAAGAAATTTGTAACCTATGACTGGCAATCGAGTGAAGAA	360
QY	361	TTTGATTTTGTAGTCTCGATTCATCATCTATGTTGAATACAGATGAGCAAGATCTT	420
Db	361	TTTGATTTTGTAGTCTCGATTCATCATCTATGTTGAATACAGATGAGCAAGATCTT	420
QY	421	TTGAGGTTGAGAGAGAGAAACCTAAGAAATTTGGATCATCAAAAAGATGCGAAGA	480
Db	421	TTGAGGTTGAGAGAGAGAAACCTAAGAAATTTGGATCATCAAAAAGATGCGAAGA	480
QY	481	AAAGTCTCAAGGTTAAATTTGGAATTAAGAAAAACATGTGGAGACAGTGACAGTGG	540
Db	481	AAAGTCTCAAGGTTAAATTTGGAATTAAGAAAAACATGTGGAGACAGTGACAGTGG	540
QY	541	ATTATCTGAGTTTCACAGTAGACCTGTGAATGCTTCATCTTCAATTCGATCGAG	600
Db	541	ATTATCTGAGTTTCACAGTAGACCTGTGAATGCTTCATCTTCAATTCGATCGAG	600
QY	601	ACTCTGTTATTAATGATACACAGATCAGAAATGTGAGGCTTTGGCGAGACGAGAAACA	660
Db	601	ACTCTGTTATTAATGATACACAGATCAGAAATGTGAGGCTTTGGCGAGACGAGAAACA	660
QY	661	GAAGAAGAAATGCGGAGACGATCATCTCCGATTCGATTCAGATGATGATATTC	720
Db	661	GAAGAAGAAATGCGGAGACGATCATCTCCGATTCGATTCAGATGATGATATTC	720
QY	721	GCTTCTCGCCTTCTGTCAACCTTTCGTCAATTCCTGTAAAGCTCATCAATTTCTGATTCG	780
Db	721	GCTTCTCGCCTTCTGTCAACCTTTCGTCAATTCCTGTAAAGCTCATCAATTTCTGATTCG	780
QY	781	CAAGTCTTATTTCTAGTTCTGATCAGATCAGACTGCGACACTTTTCTGGCTTAATTCATA	840
Db	781	CAAGTCTTATTTCTAGTTCTGATCAGATCAGACTGCGACACTTTTCTGGCTTAATTCATA	840
QY	841	GGGAGAGTGCATACCGCATCTTAAATCCTTTCCATTAAAGTGAAGTGAATGAG	900
Db	841	GGGAGAGTGCATACCGCATCTTAAATCCTTTCCATTAAAGTGAAGTGAATGAG	900
QY	901	TCTTTAATTTGAAAAAGCGATTCATGATTTCTCTGCGCAATGATGATTTCCAGATTT	960
Db	901	TCTTTAATTTGAAAAAGCGATTCATGATTTCTCTGCGCAATGATGATTTCCAGATTT	960
QY	961	TTAGGAACAAAGGTTAGATTTAAAGCTCAAAATATATTTAGTGAATTTAAAGTGAAGC	1020
Db	961	TTAGGAACAAAGGTTAGATTTAAAGCTCAAAATATATTTAGTGAATTTAAAGTGAAGC	1020
QY	1021	TCACAAGTTTCAATTTCTTATGATGATGATGATGATGATGATGATGATGATGATG	1080
Db	1021	TCACAAGTTTCAATTTCTTATGATGATGATGATGATGATGATGATGATGATGATG	1080
QY	1081	TTACTCAGACCTTTGCTTCTTTTGGAGAAATGAGTTCAGAGACTTATCAACATATGCGC	1140
Db	1081	TTACTCAGACCTTTGCTTCTTTTGGAGAAATGAGTTCAGAGACTTATCAACATATGCGC	1140

QY 1141 CAAGACCACTTGGATGCTGGAAAAATATATTTGCTATGAGACATTTGAAGCTATTC 1200
 DB 1141 CAAGACCACTTGGATGCTGGAAAAATATATTTGCTATGAGACATTTGAAGCTATTC 1200
 QY 1201 TTCAGCTGCTTCCCTGGTAAAGAGTGGAGGTCCAATATCTCCAGCCGGAACCCGAC 1260
 DB 1201 TTCAGCTGCTTCCCTGGTAAAGAGTGGAGGTCCAATATCTCCAGCCGGAACCCGAC 1260
 QY 1261 CAGTTTACAAGATATGTTCAATTAAGTCTGTGTGGTGAAGTGGATGACGATGACAACTG 1320
 DB 1261 CAGTTTACAAGATATGTTCAATTAAGTCTGTGTGGTGAAGTGGATGACGATGACAACTG 1320
 QY 1321 TAGGCTCTTTCTTCCAGGCCAATGCTGGCTGCTCTTCTTCTTGTACACTGCAACCTA 1380
 DB 1321 TAGGCTCTTTCTTCCAGGCCAATGCTGGCTGCTCTTCTTCTTGTACACTGCAACCTA 1380
 QY 1381 TCTTGGCTTTGGTGGTAAGAAATGTTCCGATGACCTTGGTTGGTTAACTTATGTT 1440
 DB 1381 TCTTGGCTTTGGTGGTAAGAAATGTTCCGATGACCTTGGTTGGTTAACTTATGTT 1440
 QY 1441 TACGTATCTTAACCTGGATTAAGTGTGACCTATGTCAGTTAATTTCTGTGCTTAT 1500
 DB 1441 TACGTATCTTAACCTGGATTAAGTGTGACCTATGTCAGTTAATTTCTGTGCTTAT 1500
 QY 1501 CAGGTTTGGAACTCTTCAACCCCTCAATGCTATGATCACTTGGGTAATATTTTCCGCT 1560
 DB 1501 CAGGTTTGGAACTCTTCAACCCCTCAATGCTATGATCACTTGGGTAATATTTTCCGCT 1560
 QY 1561 ACTAATATTCGGAAGCTTCATATTTGTGTGTTGTGATCAATAAAGTAAGTGTCTAT 1620
 DB 1561 ACTAATATTCGGAAGCTTCATATTTGTGTGTTGTGATCAATAAAGTAAGTGTCTAT 1620
 QY 1621 CAGTAATGCTGTCTTCTTGTCTTGTATGCTATTTTGTAGAATTTGACCTTTTGTG 1680
 DB 1621 CAGTAATGCTGTCTTCTTGTCTTGTATGCTATTTTGTAGAATTTGACCTTTTGTG 1680
 QY 1681 TTGCTAAACTTATTTTCTTGTGACAGGCACTGTCACCTTCATCAAGTGCCT 1740
 DB 1681 TTGCTAAACTTATTTTCTTGTGACAGGCACTGTCACCTTCATCAAGTGCCT 1740
 QY 1741 GGTTCATGCTGAACCTTAATTAATGACTTCTATTTGGGTGAAGTATTTTCTGTGAT 1800
 DB 1741 GGTTCATGCTGAACCTTAATTAATGACTTCTATTTGGGTGAAGTATTTTCTGTGAT 1800
 QY 1801 CTGATCTTTTGAATTTGCTGTGGGATTTATTTGAGAACTCTGGCTTACTGCA 1860
 DB 1801 CTGATCTTTTGAATTTGCTGTGGGATTTATTTGAGAACTCTGGCTTACTGCA 1860
 QY 1861 TCTCCTCATATATTTTCTTGTGCTTCTTCTGACATTTGGAGTCAAGTTCATAAAGT 1920
 DB 1861 TCTCCTCATATATTTTCTTGTGCTTCTTCTGACATTTGGAGTCAAGTTCATAAAGT 1920
 QY 1921 TCCAAATTTAGTACCTGTAACGGAATGTTCTTACAGATATTTATCTTTTATGAT 1980
 DB 1921 TCCAAATTTAGTACCTGTAACGGAATGTTCTTACAGATATTTATCTTTTATGAT 1980
 QY 1981 TTAAGTACAGAAACAGTGTGTGGAATGTCGTGCTGGTGAAGGAAAGAAACAAGTA 2040
 DB 1981 TTAAGTACAGAAACAGTGTGTGGAATGTCGTGCTGGTGAAGGAAAGAAACAAGTA 2040
 QY 2041 TGCAGTAGTATTAACAATATCAAGTGAAGTCTTCTCTCAACATGATCTTTTGTCT 2100
 DB 2041 TGCAGTAGTATTAACAATATCAAGTGAAGTCTTCTCTCAACATGATCTTTTGTCT 2100
 QY 2101 CCAACTTATATTAACAATATCAAGTGAAGTCTTCTCTCAACATGATCTTTTGTCT 2160
 DB 2101 CCAACTTATATTAACAATATCAAGTGAAGTCTTCTCTCAACATGATCTTTTGTCT 2160
 QY 2161 TCGGTGCTTTCAGGAGTGAAGTGTACCTGCAATTTGTAAGAGTTCGACATCAAGGT 2220
 DB 2161 TCGGTGCTTTCAGGAGTGAAGTGTACCTGCAATTTGTAAGAGTTCGACATCAAGGT 2220

QY 2221 GTTACTAATTCAGATTCGGAATGATGCTTGGGCACTTCTTGCATCGATCTGAT 2280
 DB 2221 GTTACTAATTCAGATTCGGAATGATGCTTGGGCACTTCTTGCATCGATCTGAT 2280
 QY 2281 AAAACAGTATTTTCTGGATTAATCTTATGTAATTAATCTGACGACATATGCTGG 2340
 DB 2281 AAAACAGTATTTTCTGGATTAATCTTATGTAATTAATCTGACGACATATGCTGG 2340
 QY 2341 TGGTTCTTTAAGCTTATTTCAATTAAGAGCTTCACTTCTTCTGATTTGATTTCT 2400
 DB 2341 TGGTTCTTTAAGCTTATTTCAATTAAGAGCTTCACTTCTTCTGATTTGATTTCT 2400
 QY 2401 GTACTTCTTTAACCAGTATGAATTAATGGAAGTATGATTCATCTGCTGTAACA 2460
 DB 2401 GTACTTCTTTAACCAGTATGAATTAATGGAAGTATGATTCATCTGCTGTAACA 2460
 QY 2461 CCATCTGATGCTGCTGATGTCACAAATTTCTGTGGGAAGCTGTTATTTGAACA 2520
 DB 2461 CCATCTGATGCTGCTGATGTCACAAATTTCTGTGGGAAGCTGTTATTTGAACA 2520
 QY 2521 CCATGACATTCGACATGACGAGTATGCTTGTGTAACATATTAATTTATATATC 2580
 DB 2521 CCATGACATTCGACATGACGAGTATGCTTGTGTAACATATTAATTTATATATC 2580
 QY 2581 ACATCTCATATAGGCTCAAGAAAGAAAGCCAGTATGATTAACAAGTTCCTACT 2640
 DB 2581 ACATCTCATATAGGCTCAAGAAAGAAAGCCAGTATGATTAACAAGTTCCTACT 2640
 QY 2641 GTTGGCTATTAATTTCTTGGCACTTCTGTTATATGAGCTTAAGCAGACAGATTTAT 2700
 DB 2641 GTTGGCTATTAATTTCTTGGCACTTCTGTTATATGAGCTTAAGCAGACAGATTTAT 2700
 QY 2701 AGATTTGTTAGCTGCTTCAACATGAACTTAAGCTTCTTCAATGCTTATAT 2760
 DB 2701 AGATTTGTTAGCTGCTTCAACATGAACTTAAGCTTCTTCAATGCTTATAT 2760
 QY 2761 TATGATTTCTGAGCTGATCTATATATGCTGGGGTGTCTAGTGTGGTCTTCTGTC 2820
 DB 2761 TATGATTTCTGAGCTGATCTATATATGCTGGGGTGTCTAGTGTGGTCTTCTGTC 2820
 QY 2821 TACACTTCTCCAGCATTTACCTTGTGACCAACCCCTGCACTCGAATCGAGTGCAT 2880
 DB 2821 TACACTTCTCCAGCATTTACCTTGTGACCAACCCCTGCACTCGAATCGAGTGCAT 2880
 QY 2881 AATTTGAAATTTGCTAAGAAAGAAATGCTTGTGTATATCAATGAGAACTT 2940
 DB 2881 AATTTGAAATTTGCTAAGAAAGAAATGCTTGTGTATATCAATGAGAACTT 2940
 QY 2941 TTAGATACCACTTTTATACGAAAAAAACAAGTAAGTGAAGTGAAGTGAAGT 3000
 DB 2941 TTAGATACCACTTTTATACGAAAAAAACAAGTAAGTGAAGTGAAGTGAAGT 3000
 QY 3001 CCAATTTCTAGGAATTTGACTTACTCTTGTATTTGAATTTCTGTGATCTCATACA 3060
 DB 3001 CCAATTTCTAGGAATTTGACTTACTCTTGTATTTGAATTTCTGTGATCTCATACA 3060
 QY 3061 AACACAGGAGAGAGTAAAGTATGATTTGTGATCAAACTCTGTTGCAATATATGCTC 3120
 DB 3061 AACACAGGAGAGAGTAAAGTATGATTTGTGATCAAACTCTGTTGCAATATATGCTC 3120
 QY 3121 ATGATTTAGTGGCCAGATATCAAGTTTTCTAGTAGTAAGAAAGCTGTCATGAAG 3180
 DB 3121 ATGATTTAGTGGCCAGATATCAAGTTTTCTAGTAGTAAGAAAGCTGTCATGAAG 3180
 QY 3181 ACCACTGATGCTAGATATATGGGACATATGCTCTATCTGCTGATCAAGCTTCACT 3240
 DB 3181 ACCACTGATGCTAGATATATGGGACATATGCTCTATCTGCTGATCAAGCTTCACT 3240
 QY 3241 ATCCGCTTTGTTGCTGCAATATATCAATCTGCTGAGAGAAATTTCTGCAATTA 3300
 DB 3241 ATCCGCTTTGTTGCTGCAATATATCAATCTGCTGAGAGAAATTTCTGCAATTA 3300
 QY 3301 ATTAATATGACTGTGATAGACAAAGGCAAGTTCAGAGAGACAAACGGGAAATGTT 3360

Db 3301 ATAACTATGCTGTGATAGCAAGCAAGCAGAGTTCAGAGAGACAAAGGGAATGTTG 3360
 QY 3361 GTTGGGGAAGAGCCCGTCAAGGTGATCATATTAAGCCCACTAGCTTAAGTCAAC 3420
 Db 3361 GTTGGGGAAGAGCCCGTCAAGGTGATCATATTAAGCCCACTAGCTTAAGTCAAC 3420
 QY 3421 TTACCTATATTCCTTTTCAGACCTTAACATTTTGTATGATGATCTCTTGCTCT 3480
 Db 3421 TTACCTATATTCCTTTTCAGACCTTAACATTTTGTATGATGATCTCTTGCTCT 3480
 QY 3481 CGGTTTGTGAGATTTGTGGCTGCTGATPACTACAACTGTGTGAACCTAAACTAGCT 3540
 Db 3481 CGGTTTGTGAGATTTGTGGCTGCTGATPACTACAACTGTGTGAACCTAAACTAGCT 3540
 QY 3541 TCTCTTAACGCTGAGATGAGAGCTATTTCTACATTTTACAGTTTACACCTAAATAAG 3600
 Db 3541 TCTCTTAACGCTGAGATGAGAGCTATTTCTACATTTTACAGTTTACACCTAAATAAG 3600
 QY 3601 ACCTTGTTCATCATCAATGCGCTTGAATGTGACAGTGGGATTTGGCTGCTCATTTCCAT 3660
 Db 3601 ACCTTGTTCATCATCAATGCGCTTGAATGTGACAGTGGGATTTGGCTGCTCATTTCCAT 3660
 QY 3661 ATGTTCTGAGATCTTAAGTGTCTTCTTGAGACCGCTCTCTTGATTAAGCTAA 3720
 Db 3661 ATGTTCTGAGATCTTAAGTGTCTTCTTGAGACCGCTCTCTTGATTAAGCTAA 3720
 QY 3721 AATCTATCAATCTATTTGCTATTTGAGAGCTCTTCTGCAAGTGAAGCAAGCTTAACA 3780
 Db 3721 AATCTATCAATCTATTTGCTATTTGAGAGCTCTTCTGCAAGTGAAGCAAGCTTAACA 3780
 QY 3781 TATAATCATCTTGTGTTAAACAGTCTTGACATCTTCTAGCTCATTTTCTCACCT 3840
 Db 3781 TATAATCATCTTGTGTTAAACAGTCTTGACATCTTCTAGCTCATTTTCTCACCT 3840
 QY 3841 TCTTCTCTTGTGATGACCCAGAGAGACGATGACCGATGCCGATCAAAAGTAAGCACAA 3900
 Db 3841 TCTTCTCTTGTGATGACCCAGAGAGACGATGACCGATGCCGATCAAAAGTAAGCACAA 3900
 QY 3901 TCTAATGATTTATTTTACACATCTTGTGATGATCTTCTGATGATGCTTCTACATATTCG 3960
 Db 3901 TCTAATGATTTATTTTACACATCTTGTGATGATCTTCTGATGATGCTTCTACATATTCG 3960
 QY 3961 GGTGTTGTTTGTGTTAAAGGATGAGAAATTTGAGAGCTGATTTGAG -AAGTCAAA 4019
 Db 3961 GGTGTTGTTTGTGTTAAAGGATGAGAAATTTGAGAGCTGATTTGAG -AAGTCAAA 4019
 QY 4020 TACAGATCATTCGCGGAATTTATGATGATTAACGAAGTGTGTTCTCATTTTCTACT 4079
 Db 4020 TACAGATCATTCGCGGAATTTATGATGATTAACGAAGTGTGTTCTCATTTTCTACT 4079
 QY 4080 TATTAGCTTAATGAGAGCTTGAATGATCAAAAGACCGTCCCAAAACAAATGCAAT 4139
 Db 4081 TATTAGCTTAATGAGAGCTTGAATGATCAAAAGACCGTCCCAAAACAAATGCAAT 4140
 QY 4140 TGATGATGAGAGATCTTGTGATTTGATGATGATGATGATGATGATGATGATGATGAT 4199
 Db 4141 TGATGATGAGAGATCTTGTGATTTGATGATGATGATGATGATGATGATGATGATGAT 4200
 QY 4200 GCTTATGATTTCCCTAGATTAATCAAGGATGCGCTGCTTCTAATTTTATTTATTTGA 4259
 Db 4201 GCTTATGATTTCCCTAGATTAATCAAGGATGCGCTGCTTCTAATTTTATTTATTTGA 4260
 QY 4260 TATTGATTAATGCTGCTTCTTAATGATCAAAATGATGATGATGATGATGATGATGATGAT 4319
 Db 4261 TATTGATTAATGCTGCTTCTTAATGATCAAAATGATGATGATGATGATGATGATGATGAT 4320
 QY 4320 ACTACATCAATGAGAGCTTTTATTAAGTACAGGACATTAATGATGATGATGATGATGAT 4379
 Db 4321 ACTACATCAATGAGAGCTTTTATTAAGTACAGGACATTAATGATGATGATGATGATGAT 4380
 QY 4380 TACATGATTAATCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4439
 Db 4380 TACATGATTAATCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4439

Db 4381 TACATGATTAATCAAAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 4440
 QY 4440 ACCGATTAATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4499
 Db 4441 ACCGATTAATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4500
 QY 4500 AACTATTAATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4559
 Db 4501 AACTATTAATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4560
 QY 4560 AACTATTAATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4619
 Db 4561 AACTATTAATCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4620
 QY 4620 AGCTAGAGAGCAATGCTCAAAACAAAGATTAAGAGGAGAGCTCAAGAGCTCAAC 4679
 Db 4621 AGCTAGAGAGCAATGCTCAAAACAAAGATTAAGAGGAGAGCTCAAGAGCTCAAC 4680
 QY 4680 CGAGTAGAGAGCAATGCTCAAAACAAAGATTAAGAGGAGAGCTCAAGAGCTCAAC 4739
 Db 4681 CGAGTAGAGAGCAATGCTCAAAACAAAGATTAAGAGGAGAGCTCAAGAGCTCAAC 4740
 QY 4740 GTACCAAGCTCTTCAAGCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 4799
 Db 4741 GTACCAAGCTCTTCAAGCAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 4800
 QY 4800 AAGTTTTTAAATCAAAACAAAGATTAAGAGGAGAGCTCAAGAGCTCAAC 4859
 Db 4801 AAGTTTTTAAATCAAAACAAAGATTAAGAGGAGAGCTCAAGAGCTCAAC 4860
 QY 4860 ATCTAAACAAATTTACTG 4879
 Db 4861 ATCTAAACAAATTTACTG 4880

RESULT 7
 US-09-817-774-40
 ; Sequence 40, Application US/09817774
 ; Patent No. US2002012011A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHOE, Sunghwa
 ; APPLICANT: FELDMANN A., Kenneth
 ; TITLE OF INVENTION: Dwf5 MUTANTS
 ; FILE REFERENCE: 2225-0020 / 91020.002
 ; CURRENT APPLICATION NUMBER: US/09/817,774
 ; PRIOR APPLICATION NUMBER: 60/192,202
 ; PRIOR FILING DATE: 2000-03-27
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 40
 ; LENGTH: 4880
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (671)..(757)
 ; OTHER INFORMATION: dwf5-5
 ; NAME/KEY: CDS
 ; LOCATION: (1050)..(1271)
 ; OTHER INFORMATION: dwf5-5
 ; NAME/KEY: CDS
 ; LOCATION: (1340)..(1396)
 ; OTHER INFORMATION: dwf5-5
 ; NAME/KEY: CDS
 ; LOCATION: (1505)..(1606)
 ; OTHER INFORMATION: dwf5-5
 ; NAME/KEY: CDS
 ; LOCATION: (1711)..(1776)
 ; OTHER INFORMATION: dwf5-5
 ; NAME/KEY: CDS
 ; LOCATION: (2174)..(2287)
 ; OTHER INFORMATION: dwf5-5

NAME/KEY: CDS
 LOCATION: (2418) .. (2546)
 OTHER INFORMATION: dwf5-5
 NAME/KEY: CDS
 LOCATION: (2776) .. (2874)
 OTHER INFORMATION: dwf5-5
 NAME/KEY: CDS
 LOCATION: (3259) .. (3384)
 OTHER INFORMATION: dwf5-5
 NAME/KEY: CDS
 LOCATION: (3494) .. (3559)
 OTHER INFORMATION: dwf5-5
 NAME/KEY: CDS
 LOCATION: (3636) .. (3716)
 OTHER INFORMATION: dwf5-5
 NAME/KEY: CDS
 LOCATION: (3806) .. (3883)
 OTHER INFORMATION: dwf5-5
 US-09-817-774-40

Query Match 99.7%; Score 4866.4; DB 10; Length 4880;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4878; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 AACTTGCCTTCTCCACACAGATCTCTTTCGGCTGCTAGCATTTGTTAAG 60
 DB 1 AACTTGCCTTCTCCACACAGATCTCTTTCGGCTGCTAGCATTTGTTAAG 60
 QY 61 ATGTGATGTCGAAATTAATCTGTTCTTCTATCATCATCATAGGACATA 120
 DB 61 ATGTGATGTCGAAATTAATCTGTTCTTCTATCATCATCATAGGACATA 120
 QY 121 ATTGCATCTCTTAACATTAATAATTTCTGAAATGAATAAACTGAAACGAAAA 180
 DB 121 ATTGCATCTCTTAACATTAATAATTTCTGAAATGAATAAACTGAAACGAAAA 180
 QY 181 GAATTCACACAGATCAGAAAACTAGATAGTTTCCGGGAGTTACCAAAACATT 240
 DB 181 GAATTCACACAGATCAGAAAACTAGATAGTTTCCGGGAGTTACCAAAACATT 240
 QY 241 AGAAAAATGATCTTACACAGATTCACCAATTTCTGGAATTTTCTCCAGCAAG 300
 DB 241 AGAAAAATGATCTTACACAGATTCACCAATTTCTGGAATTTTCTCCAGCAAG 300
 QY 301 CCATTTGCTGAGTAGAAGAAATTTGTAACATGACGCAATCAGAGTGAAGAA 360
 DB 301 CCATTTGCTGAGTAGAAGAAATTTGTAACATGACGCAATCAGAGTGAAGAA 360
 QY 361 TTGATTTTGTAGTCTGATTCATCATATGTTGTAATACAGAGTGAAGATCTT 420
 DB 361 TTGATTTTGTAGTCTGATTCATCATATGTTGTAATACAGAGTGAAGATCTT 420
 QY 421 TTGAGGTTGAGAGAGAGAACTAGAGATTTGGATCATTAATAACATGCGCAAG 480
 DB 421 TTGAGGTTGAGAGAGAGAACTAGAGATTTGGATCATTAATAACATGCGCAAG 480
 QY 481 AAAGTCAAGGGTAAATTTGAATTAAGAAAAACATGAGAGAGACGTAACGCGG 540
 DB 481 AAAGTCAAGGGTAAATTTGAATTAAGAAAAACATGAGAGAGACGTAACGCGG 540
 QY 541 ATTATATCTGATTCACAGTAGACCTGTGAATTCATCTTCACTGTTAACTTGCAG 600
 DB 541 ATTATATCTGATTCACAGTAGACCTGTGAATTCATCTTCACTGTTAACTTGCAG 600
 QY 601 ACTCTGTTGATTAAGTCAACAGATCAGATCTGAGGCTTTGGCCGAGAGAGAGCA 660
 DB 601 ACTCTGTTGATTAAGTCAACAGATCAGATCTGAGGCTTTGGCCGAGAGAGAGCA 660
 QY 661 GAAGAAAGAAATGGCGAGAGCTGACATTCGAGATGTTAGTACGATCAGATGTTATC 720
 DB 661 GAAGAAAGAAATGGCGAGAGCTGACATTCGAGATGTTAGTACGATCAGATGTTATC 720
 QY 721 GCTTCGCTTCTGTGACACCTTCTGATTCCTGTAAGTCAATTAATTTGATTCG 780

DB 721 GCTTCGCTTCTGTGACACCTTCTGATTCCTGTAAGTCAATTAATTTGATTCG 780
 QY 781 CAAGCTTTATTAAGTCTAGATCAGATCAGACATTTTCTGCTCTTAATTTATA 840
 DB 781 CAAGCTTTATTAAGTCTAGATCAGATCAGACATTTTCTGCTCTTAATTTATA 840
 QY 841 GCGAAGTGCATGCGCATATCTTAAATCCATTAAGTAGAGCTTGAAGAG 900
 DB 841 GCGAAGTGCATGCGCATATCTTAAATCCATTAAGTAGAGCTTGAAGAG 900
 QY 901 TCTTTAATTTGAAAAAGCATTAATGATTTCTTCTGCGCAATGATGATTTCCAGATT 960
 DB 901 TCTTTAATTTGAAAAAGCATTAATGATTTCTTCTGCGCAATGATGATTTCCAGATT 960
 QY 961 TTAGAAACCAAGGGTAGATTTAAAGCTTCAAAATCATTTAGAGTTTAACTAATGC 1020
 DB 961 TTAGAAACCAAGGGTAGATTTAAAGCTTCAAAATCATTTAGAGTTTAACTAATGC 1020
 QY 1021 TCACAAATTTGATTCCTTATGATGACATGCTACAAATGTTTCATCAGATGTTCTG 1080
 DB 1021 TCACAAATTTGATTCCTTATGATGACATGCTACAAATGTTTCATCAGATGTTCTG 1080
 QY 1081 TTACTCAGACCTTTGGCTTCTTTGGGAAATGAGTTCAAGACTTATCAATATGCG 1140
 DB 1081 TTACTCAGACCTTTGGCTTCTTTGGGAAATGAGTTCAAGACTTATCAATATGCG 1140
 QY 1141 CAAGACCCACTTTGATTTCTTGGAAATTAATTTGATAGAGCATTTGAACATATC 1200
 DB 1141 CAAGACCCACTTTGATTTCTTGGAAATTAATTTGATAGAGCATTTGAACATATC 1200
 QY 1201 TTCACCTGCTTCTGCTGTAAGAAAGTTGAGGGTCCAAATTCACACCGGAAACGAC 1260
 DB 1201 TTCACCTGCTTCTGCTGTAAGAAAGTTGAGGGTCCAAATTCACACCGGAAACGAC 1260
 QY 1261 CAGTTTAAAGGATATTTCAATTAAGTCTGTTGAGGATGATGAGATGACAACTG 1320
 DB 1261 CAGTTTAAAGGATATTTCAATTAAGTCTGTTGAGGATGATGAGATGACAACTG 1320
 QY 1321 TACGGTCTTCTTCTGACGGCCAAATGCTGCTTACTTTGACACATGACAGAACCTA 1380
 DB 1321 TACGGTCTTCTTCTGACGGCCAAATGCTGCTTACTTTGACACATGACAGAACCTA 1380
 QY 1381 TCTTGTCTTCTGCTGTAAGAAATTTCCGATGACTTTGTTGTTAATTTGTT 1440
 DB 1381 TCTTGTCTTCTGCTGTAAGAAATTTCCGATGACTTTGTTGTTAATTTGTT 1440
 QY 1441 TACGATCTTAAAGCTTGAATTAAGTTGACACTATGAGTTAATTTCTGCGTTATT 1500
 DB 1441 TACGATCTTAAAGCTTGAATTAAGTTGACACTATGAGTTAATTTCTGCGTTATT 1500
 QY 1501 CAGGTTTGAATCTTCAACCCCTGCAATTTGCTATGATACCTTGGTGAATTTTGGCC 1560
 DB 1501 CAGGTTTGAATCTTCAACCCCTGCAATTTGCTATGATACCTTGGTGAATTTTGGCC 1560
 QY 1561 ACTAATATTCGAACTTCAATTTTGTGTTTGTGTTGTTGTTAATAAGTATGCTATG 1620
 DB 1561 ACTAATATTCGAACTTCAATTTTGTGTTTGTGTTTGTGTTGTTAATAAGTATGCTATG 1620
 QY 1621 CATGAATGCTGATTTCTTGTGTTTGTGTTTGTGTTTGTGTTGTTGTTGTTG 1680
 DB 1621 CATGAATGCTGATTTCTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTGTTG 1680
 QY 1681 TTGCTAAACCTAATTTTCTTGTGTTGAGGCGCATGTTGACCTTCAAGTACTCT 1740
 DB 1681 TTGCTAAACCTAATTTTCTTGTGTTGAGGCGCATGTTGACCTTCAAGTACTCT 1740
 QY 1741 GGTTCATGCTGTAACCTAATTAATTAATTTGATTTGAGTATTTCTGCTGATCATAT 1800
 DB 1741 GGTTCATGCTGTAACCTAATTAATTAATTTGATTTGAGTATTTCTGCTGATCATAT 1800
 QY 1801 CTCATCTTTTGAATTTGCTGTTGGAATTTATTTAGGAACCTGCTTACTAGCCA 1860

Db 1801 CTGATTCCTTTGAAATGCTGTTGGGANTTATTTGAGGAACCTGCTACTAGCCA 1860
QY 1861 TCTCCATATATATTTTGGCTCTTTCTCGACATTTGAGGTCAGTTTCACATAAGT 1920
Db 1861 TCTCCATATATATTTTGGCTCTTTCTCGACATTTGAGGTCAGTTTCACATAAGT 1920
QY 1921 TCGAATTTAGTAGACCTGTAAACGGAATGTCCTTAGAGATTTATTCCTTTTATGAT 1980
Db 1921 TCGAATTTAGTAGACCTGTAAACGGAATGTCCTTAGAGATTTATTCCTTTTATGAT 1980
QY 1981 TTAGTGCAAGAAACATAGTTTGGTGATGTCGTGCGGTGAAGGAAAGCAAGTA 2040
Db 1981 TTAGTGCAAGAAACATAGTTTGGTGATGTCGTGCGGTGAAGGAAAGCAAGTA 2040
QY 2041 TGCAGTAGTTATTTAACAATATCAATGAGATCTCTCTCCACATGTACTCTTTTGGT 2100
Db 2041 TGCAGTAGTTATTTAACAATATCAATGAGATCTCTCTCCACATGTACTCTTTTGGT 2100
QY 2101 CCAACTATATCAATTAACAATATCAATGAGATCTCTCTCCACATGTACTCTTTTGGT 2160
Db 2101 CCAACTATATCAATTAACAATATCAATGAGATCTCTCTCCACATGTACTCTTTTGGT 2160
QY 2161 TCGTGCTTTGAGGGCATGAGATGTTGACCTCCGATTTGGTAAGCTTTTACATCAAGT 2220
Db 2161 TCGTGCTTTGAGGGCATGAGATGTTGACCTCCGATTTGGTAAGCTTTTACATCAAGT 2220
QY 2221 GTTACTAATTTGAGATTCGGAATGATGCTTGGGCAATCTTTCAGATCAGTACTGAT 2280
Db 2221 GTTACTAATTTGAGATTCGGAATGATGCTTGGGCAATCTTTCAGATCAGTACTGAT 2280
QY 2281 AAAACAGGTATTTTCTGGAATTTCTTATGTAATACTACATCTGACGACATAGTCTG 2340
Db 2281 AAAACAGGTATTTTCTGGAATTTCTTATGTAATACTACATCTGACGACATAGTCTG 2340
QY 2341 TGGTTTCTTAAGCCCTTATTTCAATGAGAGCTTCACTCTTTCGATTTGATTCCT 2400
Db 2341 TGGTTTCTTAAGCCCTTATTTCAATGAGAGCTTCACTCTTTCGATTTGATTCCT 2400
QY 2401 GTACTTCTTTAACAAGTATGAAATTAATGCAAAAGTATCTGATTAATGCTGTAACA 2460
Db 2401 GTACTTCTTTAACAAGTATGAAATTAATGCAAAAGTATCTGATTAATGCTGTAACA 2460
QY 2461 CCAATCTGATGCTGCTGATATGTCACAAAATCTTCTGTTGGGAGAGTGTATTTGGAACA 2520
Db 2461 CCAATCTGATGCTGCTGATATGTCACAAAATCTTCTGTTGGGAGAGTGTATTTGGAACA 2520
QY 2521 CCAATGCAATTTGCAATGACGAGGATGCTTCTGATATATATATATATATATATATAT 2580
Db 2521 CCAATGCAATTTGCAATGACGAGGATGCTTCTGATATATATATATATATATATATAT 2580
QY 2581 ACATCTCATATAGGATGACAGGAAAGAAAGCCAGTATGTAATAAAGTTTCTCTACT 2640
Db 2581 ACATCTCATATAGGATGACAGGAAAGAAAGCCAGTATGTAATAAAGTTTCTCTACT 2640
QY 2641 GTTTGGCTATATTTCTTGGCAGATTCCTGTTTATGAGCTAAGGACGACGACATATATTC 2700
Db 2641 GTTTGGCTATATATTTCTTGGCAGATTCCTGTTTATGAGCTAAGGACGACGACATATATTC 2700
QY 2701 AGATTTATGTTACTGCTTTAACAACATGAACTTAAGCTTACTTTTAAATGCTTTATAT 2760
Db 2701 AGATTTATGTTACTGCTTTAACAACATGAACTTAAGCTTACTTTTAAATGCTTTATAT 2760
QY 2761 TATGATTTCTCAGCTGATTTCTATATATGCTGGGCTTGTCTAATGCTGGCTTCTGTC 2820
Db 2761 TATGATTTCTCAGCTGATTTCTATATATGCTGGGCTTGTCTAATGCTGGCTTCTGTC 2820
QY 2821 TACACTTCTCAGGATGTAACCTTGTGAACACACCCGCGAATCTGGAATCAGGATGAT 2880
Db 2821 TACACTTCTCAGGATGTAACCTTGTGAACACACCCGCGAATCTGGAATCAGGATGAT 2880
QY 2881 AATTTGAATTTGCTAGAGAAAGAAATGCTTGTGTTAAATCAATGAGAACTT 2940
Db 2881 AATTTGAATTTGCTAGAGAAAGAAATGCTTGTGTTAAATCAATGAGAACTT 2940

QY 2941 TTAGATACACCTTTTATAGAAAAAACAAGTAGGTGACCTAATATGAAAAATC 3000
Db 2941 TTAGATACACCTTTTATAGAAAAAACAAGTAGGTGACCTAATATGAAAAATC 3000
QY 3001 CCAATTTCTAGGAATTTGCACTTAACCTTGTATTTGAATTTACTGTATCTTCCATGACA 3060
Db 3001 CCAATTTCTAGGAATTTGCACTTAACCTTGTATTTGAATTTACTGTATCTTCCATGACA 3060
QY 3061 AACACGCGGAGAGAGTAAGTATGATGTAATCAAAATGCTGTCATATATATGCTC 3120
Db 3061 AACACGCGGAGAGAGTAAGTATGATGTAATCAAAATGCTGTCATATATATGCTC 3120
QY 3121 ATGATTTGATTTGCAAGATATGATTTTCTGATGTAAGAAAGCTTTCATATAAAC 3180
Db 3121 ATGATTTGATTTGCAAGATATGATTTTCTGATGTAAGAAAGCTTTCATATAAAC 3180
QY 3181 ACCACTGATGCTACGATCATCGGACAAATATGCTATCTCCGTACCTTAAGACTTCAAT 3240
Db 3181 ACCACTGATGCTACGATCATCGGACAAATATGCTATCTCCGTACCTTAAGACTTCAAT 3240
QY 3241 ATCCGCTTTGCTGAGTTGCAATATATACATTTCTGTTGAGAAATCTGATTTAC 3300
Db 3241 ATCCGCTTTGCTGAGTTGCAATATATACATTTCTGTTGAGAAATCTGATTTAC 3300
QY 3301 ATAACTATGATGCTGATGTAAGGCAAGAGCTCAGAGGACAAAGGGAAATGTTG 3360
Db 3301 ATAACTATGATGCTGATGTAAGGCAAGAGCTCAGAGGACAAAGGGAAATGTTG 3360
QY 3361 GTTTGGGGAAGACCCCGTCAAAAGGTGATCATATATATAGCCCAATAGCTTAGTGAAC 3420
Db 3361 GTTTGGGGAAGACCCCGTCAAAAGGTGATCATATATATAGCCCAATAGCTTAGTGAAC 3420
QY 3421 TTAGCTATATTTCTTTTACAGCTTAACATTTTGAATGATGAATCTCTGCTCTCT 3480
Db 3421 TTAGCTATATTTCTTTTACAGCTTAACATTTTGAATGATGAATCTCTGCTCTCT 3480
QY 3481 CGTTTGTGCAATTTGCGGCTGATATCAATCAACATCTGTAACATAAAGTACTAGCT 3540
Db 3481 CGTTTGTGCAATTTGCGGCTGATATCAATCAACATCTGTAACATAAAGTACTAGCT 3540
QY 3541 TCTCTTAACGCTGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
Db 3541 TCTCTTAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3600
QY 3601 ACCTGCTTATCATCAATGAGCTTTGATGATGATGATGATGATGATGATGATGATGAT 3660
Db 3601 ACCTGCTTATCATCAATGAGCTTTGATGATGATGATGATGATGATGATGATGATGAT 3660
QY 3661 ATGTCCTGAGATCTTAAGTGTCTTCTGACCGCTACCGGCTCTCTGATTAAGTAA 3720
Db 3661 ATGTCCTGAGATCTTAAGTGTCTTCTGACCGCTACCGGCTCTCTGATTAAGTAA 3720
QY 3721 AATACATCATCTATTTGCTATATGAAAGCTCTTCTGAAAGTAAAGCAGCTTAACA 3780
Db 3721 AATACATCATCTATTTGCTATATGAAAGCTCTTCTGAAAGTAAAGCAGCTTAACA 3780
QY 3781 TATTAATCATCTTTGTTTAAACCACTTCTGCAATCTTCTGCAATATTTCTACCT 3840
Db 3781 TATTAATCATCTTTGTTTAAACCACTTCTGCAATCTTCTGCAATATTTCTACCT 3840
QY 3841 TCTTCTCTTGTGAGGACGAGAGAGAGAGATGACGATGACGATGACGATGACGATGAC 3900
Db 3841 TCTTCTCTTGTGAGGACGAGAGAGAGAGATGACGATGACGATGACGATGACGATGAC 3900
QY 3901 TCTAATGATTTATTTTACATCTCTTGTATCTTACTGATTTGCTGCAATATTTG 3960
Db 3901 TCTAATGATTTATTTTACATCTCTTGTATCTTACTGATTTGCTGCAATATTTG 3960
QY 3961 GGTTTTGTGTTGTAAGAAAGTATGGAATATTTGGAAGCTGTATTTGAG - AAGTCAAA 4019
Db 3961 GGTTTTGTGTTGTAAGAAAGTATGGAATATTTGGAAGCTGTATTTGAG - AAGTCAAA 4020

QY 361 TTTGATTTTGTAGTCTGATTCATCACTATGTGTGATACAGAGATGACGAGATCTT 420
|||||
Db 361 TTTGATTTTGTAGTCTGATTCATCACTATGTGTGATACAGAGATGACGAGATCTT 420
QY 421 TTTGAGGTTGAG 480
|||||
Db 421 TTTGAGGTTGAG 480
QY 481 AAAGCTCAAGGTTAAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAA 540
|||||
Db 481 AAAGCTCAAGGTTAAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAAATTTGAA 540
QY 541 ATTTATCTGAGTTTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 600
|||||
Db 541 ATTTATCTGAGTTTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 600
QY 601 ACTCTGTTGATTGATCAACACAGATCAAGATCTGAGGCTTGGCCGAGAGAGAGAGA 660
|||||
Db 601 ACTCTGTTGATTGATCAACACAGATCAAGATCTGAGGCTTGGCCGAGAGAGAGAGA 660
QY 661 GAAG 720
|||||
Db 661 GAAG 720
QY 721 GCTTCTGAGCTTCTGTCACCTTCTGTCATCTCTGTAAGCTTCATTTTCTGATTCG 780
|||||
Db 721 GCTTCTGAGCTTCTGTCACCTTCTGTCATCTCTGTAAGCTTCATTTTCTGATTCG 780
QY 781 CAAGCTTTATTTGATTTCTGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 840
|||||
Db 781 CAAGCTTTATTTGATTTCTGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG 840
QY 841 GCGAGAGTGCATGCGAGATTCCTTAATCTCTTCCATTTAGTACGATTCATTAAGAG 900
|||||
Db 841 GCGAGAGTGCATGCGAGATTCCTTAATCTCTTCCATTTAGTACGATTCATTAAGAG 900
QY 901 TCTTTAATGAAAAAGCATTCATGATTTCTCTGCGCAATGTGAGTTTCCAGATTT 960
|||||
Db 901 TCTTTAATGAAAAAGCATTCATGATTTCTCTGCGCAATGTGAGTTTCCAGATTT 960
QY 961 TTAGGAAACAGAGGTAGATTTAAAGTACAAATCATATTTAGTACGATTTTAAATAGC 1020
|||||
Db 961 TTAGGAAACAGAGGTAGATTTAAAGTACAAATCATATTTAGTACGATTTTAAATAGC 1020
QY 1021 TCACAGTTTCATTTCTTTATGATGAGATGATACAAATGATTCATCAGAGATGATTCG 1080
|||||
Db 1021 TCACAGTTTCATTTCTTTATGATGAGATGATACAAATGATTCATCAGAGATGATTCG 1080
QY 1081 TTTACTGAGCTTTGCTTTTGGGAGATGAGTTCAAGGACTTATCAATATAGC 1140
|||||
Db 1081 TTTACTGAGCTTTGCTTTTGGGAGATGAGTTCAAGGACTTATCAATATAGC 1140
QY 1141 CAAGAGCCACTTTGATTTGCTGGAATTTATTTTCTATGAGCATTTGAAGCTATTC 1200
|||||
Db 1141 CAAGAGCCACTTTGATTTGCTGGAATTTATTTTCTATGAGCATTTGAAGCTATTC 1200
QY 1201 TTTGAGTCTTCTGCTGCTGTAAGAGATGAGGTTCCAAATCTCCAGCCGGAACCGAG 1260
|||||
Db 1201 TTTGAGTCTTCTGCTGCTGTAAGAGATGAGGTTCCAAATCTCCAGCCGGAACCGAG 1260
QY 1261 CAGTTTACAGGATGATTTCAATTTAGTCTTGTGAGGAGATGATGACAGATGACAACTG 1320
|||||
Db 1261 CAGTTTACAGGATGATTTCAATTTAGTCTTGTGAGGAGATGATGACAGATGACAACTG 1320
QY 1321 TACGCTCTTTCTGAGGCAATGCTGCTGCTTACTTTGATGACACTTACAACTATTC 1380
|||||
Db 1321 TACGCTCTTTCTGAGGCAATGCTGCTGCTTACTTTGATGACACTTACAACTATTC 1380
QY 1381 TCTTGTCTTTGTGATGAGAGATTTTCCGATGACTTTGTTGTTTAAATTTATTTGTT 1440
|||||
Db 1381 TCTTGTCTTTGTGATGAGAGATTTTCCGATGACTTTGTTGTTTAAATTTATTTGTT 1440

QY 1441 TACGATCTTACGCTTGGATTAAGTGTGACCTATGATGATTAATTTCTGCTGATTT 1500
|||||
Db 1441 TACGATCTTACGCTTGGATTAAGTGTGACCTATGATGATTAATTTCTGCTGATTT 1500
QY 1501 CAGTTTGGATTTCTACACCTTGCATTTGATGATGATGATGATGATGATGATGATGATG 1560
|||||
Db 1501 CAGTTTGGATTTCTACACCTTGCATTTGATGATGATGATGATGATGATGATGATGATG 1560
QY 1561 ACTAATTTGGAAGCTTCATATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1620
|||||
Db 1561 ACTAATTTGGAAGCTTCATATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1620
QY 1621 CATGAGGAGTGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680
|||||
Db 1621 CATGAGGAGTGTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1680
QY 1681 TTGCTAAACTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
|||||
Db 1681 TTGCTAAACTTATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1740
QY 1741 GCTTCAATGTTGATACCTAATTAATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
|||||
Db 1741 GCTTCAATGTTGATACCTAATTAATGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1800
QY 1801 CTGATCTTTTGAATTTGCTTGTGATTTATTTGAGAGACTCTGCTTCTTCTTCTTCTTCT 1860
|||||
Db 1801 CTGATCTTTTGAATTTGCTTGTGATTTATTTGAGAGACTCTGCTTCTTCTTCTTCTTCT 1860
QY 1861 TCTCTCATTAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1920
|||||
Db 1861 TCTCTCATTAATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1920
QY 1921 TCCATTTTATGATGACCTGTAACGGAATGTTCTTACGATTAATTTATTTATTTAT 1980
|||||
Db 1921 TCCATTTTATGATGACCTGTAACGGAATGTTCTTACGATTAATTTATTTATTTATTTAT 1980
QY 1981 TTAAGTACAGAAAGTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2040
|||||
Db 1981 TTAAGTACAGAAAGTATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2040
QY 2041 TGCAGTATTTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2100
|||||
Db 2041 TGCAGTATTTATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2100
QY 2101 CCAACTTATCATTTTACCAATCCAAATTTGTCGAATCTTTCTCACTGATTTTCTGCG 2160
|||||
Db 2101 CCAACTTATCATTTTACCAATCCAAATTTGTCGAATCTTTCTCACTGATTTTCTGCG 2160
QY 2161 TCGTGTGTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
|||||
Db 2161 TCGTGTGTCAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
QY 2221 GTTTACTAATTTGCAATTCGAAATGATGATGATGATGATGATGATGATGATGATGAT 2280
|||||
Db 2221 GTTTACTAATTTGCAATTCGAAATGATGATGATGATGATGATGATGATGATGATGAT 2280
QY 2281 AAAACAGTATTTTCTGATTTCTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2340
|||||
Db 2281 AAAACAGTATTTTCTGATTTCTTATTTATTTATTTATTTATTTATTTATTTATTTAT 2340
QY 2341 TGGTTCTTTAAGTCTTATTTCTTATTTAGAGACTTCACTCTTCTTCTTCTTCTTCTTCT 2400
|||||
Db 2341 TGGTTCTTTAAGTCTTATTTCTTATTTAGAGACTTCACTCTTCTTCTTCTTCTTCTTCT 2400
QY 2401 GTACTTTCTTTAACCAGATGAAATTAATGCAAGATGATGATGATGATGATGATGATG 2460
|||||
Db 2401 GTACTTTCTTTAACCAGATGAAATTAATGCAAGATGATGATGATGATGATGATGATGATG 2460
QY 2461 CCATCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
|||||
Db 2461 CCATCTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
QY 2521 CCATGAGATGCAATGACGAGGATGCTTCTGATGATGATGATGATGATGATGATGATG 2580
|||||

D	2521	CCATGGACATTGCACATGACCGGAGGTGCTCTGGTACATATATAAATTTATATC	2580
Q	2581	ACATCTTCATATAGGGATACAAAGAAAGCCAGTAGTATATACAAAGTTCTCTACT	2640
D	2581	ACATCTTCATATAGGGATACAAAGAAAGCCAGTAGTATATACAAAGTTCTCTACT	2640
Q	2641	GTTTGGCATATATTTCCCTGGCAGTCTGTTTATATGAGCTACAGCACACATATATTC	2700
D	2641	GTTTGGCATATATTTCCCTGGCAGTCTGTTTATATGAGCTACAGCACACATATATTC	2700
Q	2701	AGATTATGTACTGCGCTTTAACACATGAACCTTAAGCTTACTTTCAATGCTTTATAT	2760
D	2701	AGATTATGTACTGCGCTTTAACACATGAACCTTACTTTCAATGCTTTATATAT	2760
Q	2761	TATGATTTCTTAGCTGGATTTCTATATATGCGGGGTGCTAGTGGGGCTCTGTC	2820
D	2761	TATGATTTCTTAGCTGGATTTCTATATATGCGGGGTGCTAGTGGGGCTCTCTGTC	2820
Q	2821	TACACTTTCCAGCGATTTACCTTGTGAAACCCCGTCGAACTCGAACTCAGGTGCAT	2880
D	2821	TACACTTTCTCAGCGATTTACCTTGTGAAACCCCGTCGAACTCGAACTCAGGTGCAT	2880
Q	2881	AATTTTGAATTTGTCTAGAGAAAGAAATGCTCTTTGTGTATATCCATGAGAACTT	2940
D	2881	AATTTTGAATTTGTCTAGAGAAAGAAATGCTCTTTGTGTATATCCATGAGAACTT	2940
Q	2941	TTAGATATCCACTTTTATATACGAAAAAACAAGTAGGTGAGCTAAATGTGAAATC	3000
D	2941	TTAGATATCCACTTTTATATACGAAAAAACAAGTAGGTGAGCTAAATGTGAAATC	3000
Q	3001	CCAAATTTCTAGGAATTTGACATTAACCTTGTATTGAATTACTTGGATCTTCCATTA	3060
D	3001	CCAAATTTCTAGGAATTTGACATTAACCTTGTATTGAATTACTTGGATCTTCCATTA	3060
Q	3061	AACACAGGGAGAGGTAAAGTAGATTGTTGATCAAATCCTCGTGGCAATATATCTC	3120
D	3061	AACACAGGGAGAGAGTAAAGTAGATTGTTGATCAAATCCTCGTGGCAATATATCTC	3120
Q	3121	ATGATTTAGTTGCCAGATATCAGTTTTCTAGTAGTAGAAGAAAGCTTTCATGAAAC	3180
D	3121	ATGATTTAGTTGCCAGATATCAGTTTTCTAGTAGTAGAAGAAAGCTTTCATGAAAC	3180
Q	3181	ACCAGTATGCTACAGCATATGCGGACATATGCGATCTCGTACCTCAAGCTCAGTAT	3240
D	3181	ACCAGTATGCTACAGCATATGCGGACATATGCGATCTCGTACCTCAAGCTCAGTAT	3240
Q	3241	ATCCGCTTTTGGTTCACATTTGACATATCAATTCGTTGAGAGAAATTCGTGCATTTAC	3300
D	3241	ATCCGCTTTTGGTTCACATTTGACATATCAATTCGTTGAGAGAAATTCGTGCATTTAC	3300
Q	3301	ATTAACATATGACTGTGATAGACAAAGGCAAGAGTTCAGAGAGCAAAAGGGAATGTTG	3360
D	3301	ATTAACATATGACTGTGATAGACAAAGGCAAGAGTTCAGAGAGCAAAAGGGAATGTTG	3360
Q	3361	GTTTGGGGAAGAGCCCGTCAAGAGGTGTGATCATATAAGCCCACTAGCTTAGTGAC	3420
D	3361	GTTTGGGGAAGAGCCCGTCAAGAGGTGTGATCATATAAGCCCACTAGCTTAGTGAC	3420
Q	3421	TTAGCATATTTCTTTTGACATTAACATTTTGTATGTATGTAATCTCTTGCTGCT	3480
D	3421	TTAGCATATTTCTTTTGACATTAACATTTTGTATGTATGTAATCTCTTGCTGCT	3480
Q	3481	CGGTTTGTTCAGATTTGTGCGTCTATATCAACAACATCTGGTGAACCTTAACCTAGCT	3540
D	3481	CGGTTTGTTCAGATTTGTGCGTCTATATCAACAACATCTGGTGAACCTTAACCTAGCT	3540
Q	3541	TCTCTTAAAGCTGTGAGTGTGAGACTATTTACATTTACAGTTACAACTTAATAATAG	3600
D	3541	TCTCTTAAAGCTGTGAGTGTGAGACTATTTACATTTACAGTTACAACTTAATAATAG	3600
Q	3601	ACGTTGTTCAATCAATGAGCTTTGAATGTGACAGTGGGATTTGGCTGTCAATTTCCATT	3660
D	3601	ACGTTGTTCAATCAATGAGCTTTGAATGTGACAGTGGGATTTGGCTGTCAATTTCCATT	3660
D	3601	ACGTTGTTCAATCAATGAGCTTTGAATGTGACAGTGGGATTTGGCTGTCAATTTCCATT	3660
Q	3661	ATGTTCCGAGATCTTAAGTCTCTTCTTTGAGACCGTACCGGCGCTCTTGATTAAGTAA	3720
D	3661	ATGTTCCGAGATCTTAAGTCTCTTCTTTGAGACCGTACCGGCGCTCTTGATTAAGTAA	3720
Q	3721	AATACATATCATTTATTTGCTATATTTGGAACCTCTTCTGCAAGTGAAGACACTTAACA	3780
D	3721	AATACATATCATTTATTTGCTATATTTGGAACCTCTTCTGCAAGTGAAGACACTTAACA	3780
Q	3781	TATAATCATCTTTGTTTAAACCACTTCTGGCATACTTCTAGTGCATATTTCTCACCTT	3840
D	3781	TATAATCATCTTTGTTTAAACCACTTCTGGCATACTTCTAGTGCATATTTCTCACCTT	3840
Q	3841	TCTCTCTTTATGAGCCCAAGAGAGAGATGACCGATGCGATCAAAAGTAAACACCAA	3900
D	3841	TCTCTCTTTATGAGCCCAAGAGAGAGATGACCGATGCGATCAAAATTAAGACCAA	3900
Q	3901	TCTAAATCATTTATTTTACACATCTCTTTGATCTTACTGTGATTTGCTACAAATATCG	3960
D	3901	TCTAAATCATTTATTTTACACATCTCTTTGATCTTACTGTGATTTGCTACAAATATCG	3960
Q	3961	GTTTTGTTTTGGTAAAAAGGTATGGGAAATTTGGAAGCTGTATTTGTGAG -AAGTCAAA	4019
D	3961	GTTTTGTTTTGGTAAAAAGGTATGGGAAATTTGGAAGCTGTATTTGTGAGAAAGTCAA	4020
Q	4020	TTACAGATCATCTCCGGAATTTATTTGATTGTAAAGAGTGTGTTCTCATTTTCTACT	4079
D	4021	TTACAGATCATCTCCGGAATTTATTTGATTGTAAAGAGTGTGTTCTCATTTTCTACT	4080
Q	4080	TTATTAAGTTAATGCAAGTGGGAATCATCAAAAGACCGTCCCAAAACAAAATATGCAAT	4139
D	4081	TTATTAAGTTAATGCAAGTGGGAATCATCAAAAGACCGTCCCAAAACAAAATATGCAAT	4140
Q	4140	TGATGAGATAGACATCTTTTGTGCTGATTTGTATGCTATAGGTTTCAATCTCAGTAC	4199
D	4141	TGATGAGATAGACATCTTTTGTGCTGATTTGTATGCTATAGGTTTCAAACTCAGTAC	4200
Q	4200	GCTTATGATTTCCCTAATATACAAATTTAGCCGCGCTTCTCAATTTATATTTGA	4259
D	4201	GCTTATGATTTCCCTAATATATCAAAATTTAGCCGCGCTTCTCAATTTATATTTGA	4260
Q	4260	TATTTTGAATTAATGCTTTTAAATGATCAATTAAGTTGATTCATCATTAATATTTT	4319
D	4261	TATTTTGAATTAATGCTTTTAAATGATCAATTAAGTTGATTCATCATTAATATTTT	4320
Q	4320	ACTACATCAATGAAGCTTTTATATAGTACGAGCACTAATTGAACCTTAATATAA	4379
D	4321	ACTACATCAATGAAGCTTTTATATAGTACGAGCACTAATTGAACCTTAATATAA	4380
Q	4380	TACATGAATTAATCAAAAGAAATTTGATAGTTCCTTCAAGTTTTCATCACAAAGACTC	4439
D	4381	TACATGAATTAATCAAAAGAAATTTGATAGTTCCTTCAAGTTTTCATCACAAAGACTC	4440
Q	4440	ACCGACATTAATCAAGTTGACTTCACTTTTGAAGCTTAATTTGTTCTGGAGATCCAA	4499
D	4441	ACCGACATTAATCAAGTTGACTTCACTTTTGAAGCTTAATTTGTTCTGGAGATCCAA	4

QY 4740 GTACCAAGCTCCTCAGCAAAATTTGTATCTATGTTGGATATCAGATCTAATCTC 4799
 |||||
 Db 4741 GTACCAAGCTCCTCAGCAAAATTTGTATCTATGTTGGATATCAGATCTAATCTC 4800
 QY 4800 AAAGTTTAAATACAAAACCTTACCATCTCTATTTGTATGTCAGCAAACTAATCTC 4859
 |||||
 Db 4801 AAAGTTTAAATACAAAACCTTACCATCTCTATTTGTATGTCAGCAAACTAATCTC 4860
 QY 4860 ATCTAAACAATTTACTG 4879
 |||||
 Db 4861 ATCTAAACAATTTACTG 4880

RESULT 9

US-09-938-842A-4796
 ; Sequence 4796, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 4796
 ; LENGTH: 543
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-4796

Query Match 11.18; Score 541.4; DB 9; Length 543;
 Best Local Similarity 99.8%; Pred. No. 3,7e-112;
 Matches 542; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 128 CTCTTAACATCTAATAATTTCTGAAATGAAAAAACTGAGAAAGCGCAAAAGATTC 187
 |||||
 Db 1 CTCTTAACATCTAATAATTTCTGAAATGAAAAAACTGAGAAAGCGCAAAAGATTC 60
 QY 188 ACACAGCATCAGAAACTAGATCTAGGTTTCGCCGGAGTTTACCAAAAACATTAGAAAA 247
 |||||
 Db 61 ACACAGCATCAGAAACTAGATCTAGGTTTCGCCGGAGTTTACCAAAAACATTAGAAAA 120
 QY 248 ATGCATCTTACAGCAATTCACCAATTTCTGAGCAATTTTCTTCCAAAGCCATTGC 307
 |||||
 Db 121 ATGCATCTTACAGCAATTCACCAATTTCTGAGCAATTTTCTTCCAAAGCCATTGC 180
 QY 308 TTGGAGTAAGAAAGTAATGTAATCTATGATCGCAATCGAGTGAAGAGTAATTTGAT 367
 |||||
 Db 181 TTGGAGTAAGAAAGTAATGTAATCTATGATCGCAATCGAGTGAAGAGTAATTTGAT 240
 QY 368 TTGTTAGTCTCGATTCATCACTATGTGTTGAATACAGAGATGAGCAAGATCTTTGAGG 427
 |||||
 Db 241 TTGTTAGTCTCGATTCATCACTATGTGTTGAATACAGAGATGAGCAAGATCTTTGAGG 300
 QY 428 TTGAGAGAGAAAGCTAAGAAAGTTGGATCAATCAAAAACGATGGCGAAGAAAGCT 487
 |||||
 Db 301 TTGAGAGAGAAAGCTAAGAAAGTTGGATCAATCAAAAACGATGGCGAAGAAAGCT 360
 QY 488 CAAGGTAATTAATGGAATTAAGAAAAACATGTGAGAGACGATGACACCGTGGATTATC 547
 |||||
 Db 361 CAAGGTAATTAATGGAATTAAGAAAAACATGTGAGAGACGATGACACCGTGGATTATC 420

QY 548 TGAGTTTCAAGTAGAGACTTGAAGTTTCATCTCTTCACTGTTAACTTGCACACTCTGT 607
 |||||
 Db 421 TGAGTTTCAAGTAGAGACTTGAAGTTTCATCTCTTCACTGTTAACTTGCACACTCTGT 480
 QY 608 TGATTTAGTCAACAGATCTGAGATCTGAGGCTTTGGCCGAGAGAGAGAAAGAGAG 667
 |||||
 Db 481 TGATTTAGTCAACAGATCTGAGATCTGAGGCTTTGGCCGAGAGAGAGAGAGAGAG 540
 QY 668 AAA 670
 |||||
 Db 541 AAA 543

RESULT 10

US-09-938-842A-3847
 ; Sequence 3847, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIPI300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 3847
 ; LENGTH: 2000
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-3847

Query Match 5.5%; Score 267.6; DB 9; Length 2000;
 Best Local Similarity 98.3%; Pred. No. 1.6e-50;
 Matches 281; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 4595 CTCACAGTGGAGATGTATACAAAAAGCTAGAGAGCAAAATCTCAAAACAAAAGACATA 4654
 |||||
 Db 1 CTCACAGTGGAGATGTATACAAAAAGCTAGAGAGCAAAATCTCAAAACAAAAGACATA 60
 QY 4655 AGAAGGGAGCTCAGAAAGCTCAACCGAGTAGAGATCAAGTTCGTAAGAACAGAAC 4714
 |||||
 Db 61 AGAAGGGAGCTCAGAAAGCTCAACCGAGTAGAGATCAAGTTCGTAAGAACAGAAC 120
 QY 4715 ATCATTTACTTGTAGGCGGCTTGTGACCAAGCTCTTACAGCAAAATTTGATTTATG 4774
 |||||
 Db 121 ATCATTTACTTGTAGGCGGCTTGTGACCAAGCTCTTACAGCAAAATTTGATTTATG 180
 QY 4775 TTTGATATCATGAGATCTTAACTTCAAAAGTTTAAATACAAAACTTACATTTCTTAA 4834
 |||||
 Db 181 TTTGATATCATGAGATCTTAACTTCAAAAGTTTAAATACAAAACTTACATTTCTTAA 240
 QY 4835 TTTGTAGTCAAAAGCT-AACAATTTCACTTAAACAATTTTACTG 4879
 |||||
 Db 241 TTTGTAGTCAAAAGCTTAAACAATTTTCACTTAAACAATTTTACTG 286

RESULT 11

US-09-817-774-44
 ; Sequence 44, Application US/09817774
 ; Patent No. US20020120111A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHOE, Sunghwa
 ; APPLICANT: FELDMANN A., Kenneth
 ; TITLE OF INVENTION: Dwf5 MUTANTS

Query Match	4.6%;	Score 224;	DB 9;	Length 1299;
Best Local Similarity	100.0%;	Pred. No. 8.2e-41;		
Matches 224;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

RESULT 14
US-09-983-965-2109
; Sequence 2109, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Ningbing
; APPLICANT: Byalt, John C.

```

RESULT 15
US-09-960-352-3400
: Sequence 3400, Application US/09960352
: Patent No. US20020137139A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Nengbing
: APPLICANT: Byatt, John C.
: APPLICANT: Mathialagan, Nagappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
: FILE REFERENCE: 16511.006/37-21(10298)C
: CURRENT APPLICATION NUMBER: US/09/960,352
: CURRENT FILING DATE: 2001-09-24
: NUMBER OF SEQ ID NOS: 15112
: SEQ ID NO 3400
: LENGTH: 446
: TYPE: DNA

```

Search completed: January 16, 2003, 21:37:18
Job time : 221 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 13:13:12 : Search time 911 Seconds
(without alignments)
12060.922 Million cell updates/sec

Title: US-09-817-774-30

Perfect score: 4879
Sequence: 1 aactgtcctcctccacaa.....atctaaacaatttactgt 4879

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: N_Geneseq_101002.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	260.4	5.3	1336	21	AAC40077
2	256.8	5.3	1501	21	AAC31344
3	237.4	4.9	1194	21	AAC36661
4	224	4.6	1496	17	AAT39358
5	87.8	1.8	354	21	AA69663
6	79.2	1.6	629	21	AAC6614
7	69.2	1.4	9539	22	AA545347
8	69.2	1.4	9539	24	ABK28180
9	68.8	1.4	6668	24	ABL33697

10	66.4	1.4	887	22	AA194064	Human neuroblastom
11	64.2	1.3	34769	22	AA546774	Tumour suppressor
12	63.8	1.3	6644	20	AA333181	Base sequence of t
13	63.8	1.3	7372	20	AA333182	Base sequence of t
14	63.8	1.3	7797	20	AA333180	Cowpox virus bar f
15	63.8	1.3	7996	20	AA333184	Base sequence of t
16	63.8	1.3	14006	24	ABL33958	Human immune syste
17	62	1.3	875	22	AA195044	Human neuroblastom
18	61.6	1.3	18598	22	ABL33286	Human immune syste
19	61.2	1.3	9964	24	ABL32098	Human immune syste
20	60.6	1.2	5413	24	ABL32564	Human immune syste
21	60.4	1.2	6668	24	ABL33696	Human immune syste
22	60.2	1.2	50000	24	ABL55643	AMEPV genome fragm
23	60	1.2	6767	22	AA546608	Tumour suppressor
24	59.6	1.2	16033	24	ABL33404	Human immune syste
25	58.4	1.2	8781	24	ABL33687	Human immune syste
26	58.4	1.2	13453	24	ABN80176	Human chemically m
27	58	1.2	18624	24	ABL33703	Human immune syste
28	57.8	1.2	819	22	AA194974	Human neuroblastom
29	57.8	1.2	7571	22	ABL32527	Human immune syste
30	57.6	1.2	6002	24	ABL32953	Human immune syste
31	57.6	1.2	6048	24	ABQ67002	Human angioogenesis
32	57.6	1.2	6314	24	ABL54314	Chemically treated
33	57.4	1.2	6161	24	ABL32411	Human immune syste
34	57.2	1.2	556	24	ABO36996	Oligonucleotide fo
35	57.2	1.2	556	24	ABO36997	Chemically pretrea
36	56.8	1.2	6032	24	AA563317	Human immune syste
37	56.8	1.2	13420	24	ABL32916	Human immune syste
38	56.6	1.2	8079	24	ABL92313	Chemically treated
39	56.6	1.2	12409	24	AA563312	Human immune syste
40	56.2	1.2	12337	24	ABL34358	Human immune syste
41	56	1.1	6898	24	ABN80222	Human chemically m
42	56	1.1	7442	22	AA546686	Tumour suppressor
43	56	1.1	7895	24	ABK28364	DNA transcription
44	55.6	1.1	5406	24	ABL33098	Human immune syste
45	55.6	1.1	5750	22	AA546708	Tumour suppressor

ALIGNMENTS

RESULT 1
AAC40077
ID AAC40077 standard; DNA; 1336 Bp.
AC AAC40077;
XX
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 26963.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-r2.
PN
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149729.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.

PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 5.3%; Score 260.4; DB 21; Length 1336;
Best Local Similarity 99.3%; Pred. No. 2.2e-50;

Matches 272; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 3978 AAGGTATGGAAATATTGGAAGCTGATTTGTGAG-AAGTCAAAATACAGATCATTCGCGG 4036
DB 1063 AAGGTATGGAAATATTGGAAGCTGATTTGTGGAAGTCAAAATACAGATCATTCGCGG 1122
QY 4037 AATTATGATTTGTAACGAAGTCTGTTCTCATTTTCTACTTATTTAGCTTAATTCGAA 4096
DB 1123 AATTATGATTTGTAACGAAGTCTGTTCTCATTTTCTACTTATTTAGCTTAATTCGAA 1182
QY 4097 CGTTGAAATCAAAAGACGCGCCAAACAAAATGCAAAATGATGATGACATTC 4156
DB 1183 CGTTGAAATCAAAAGACGCGCCAAACAAAATGCAAAATGATGACATTC 1242
QY 4157 TTTTGTGATTTGATGCTATAGTGTTCCTCAATCTAGCTAGCTATATTCCTCA 4216
DB 1243 TTTTGTGATTTGATGCTATAGTGTTCCTCAATCTAGCTAGCTATATTCCTCA 1302
QY 4217 GATTATCAAAAGTTCGCTGCGCTTTCTAATTTT 4250
DB 1303 GATTATCAAAAGTTCGCTGCGCTTTCTAATTTT 1336

RESULT 2

AAC51344
ID AAC51344 standard; DNA; 1501 BP.

AC AAC51344;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 68170.

KM Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129807.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0133484.
PR 05-MAY-1999; 99US-0133485.
PR 06-MAY-1999; 99US-0133486.
PR 07-MAY-1999; 99US-0133487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 9905-0143624.
 PR 15-JUL-1999; 9905-0144005.
 PR 16-JUL-1999; 9905-0144085.
 PR 16-JUL-1999; 9905-0144086.
 PR 19-JUL-1999; 9905-0144325.
 PR 19-JUL-1999; 9905-0144331.
 PR 19-JUL-1999; 9905-0144332.
 PR 19-JUL-1999; 9905-0144333.
 PR 19-JUL-1999; 9905-0144334.
 PR 19-JUL-1999; 9905-0144335.
 PR 20-JUL-1999; 9905-0144352.
 PR 20-JUL-1999; 9905-0144632.
 PR 20-JUL-1999; 9905-0144884.
 PR 21-JUL-1999; 9905-0144814.
 PR 21-JUL-1999; 9905-0145086.
 PR 21-JUL-1999; 9905-0145088.
 PR 22-JUL-1999; 9905-0145085.
 PR 22-JUL-1999; 9905-0145087.
 PR 22-JUL-1999; 9905-0145089.
 PR 22-JUL-1999; 9905-0145145.
 PR 23-JUL-1999; 9905-0145145.
 PR 23-JUL-1999; 9905-0145218.
 PR 23-JUL-1999; 9905-0145224.
 PR 26-JUL-1999; 9905-0145276.
 PR 27-JUL-1999; 9905-0145913.
 PR 27-JUL-1999; 9905-0145918.
 PR 27-JUL-1999; 9905-0145919.
 PR 28-JUL-1999; 9905-0145951.
 PR 02-AUG-1999; 9905-0146386.
 PR 02-AUG-1999; 9905-0146388.
 PR 02-AUG-1999; 9905-0146389.
 PR 03-AUG-1999; 9905-0147038.
 PR 04-AUG-1999; 9905-0147204.
 PR 04-AUG-1999; 9905-0147302.
 PR 05-AUG-1999; 9905-0147192.
 PR 05-AUG-1999; 9905-0147260.
 PR 06-AUG-1999; 9905-0147303.
 PR 06-AUG-1999; 9905-0147416.
 PR 09-AUG-1999; 9905-0147493.
 PR 09-AUG-1999; 9905-0147935.
 PR 10-AUG-1999; 9905-0148171.
 PR 11-AUG-1999; 9905-0148319.
 PR 12-AUG-1999; 9905-0148341.
 PR 13-AUG-1999; 9905-0148565.
 PR 13-AUG-1999; 9905-0148684.
 PR 16-AUG-1999; 9905-0149368.
 PR 17-AUG-1999; 9905-0149175.
 PR 18-AUG-1999; 9905-0149426.
 PR 20-AUG-1999; 9905-0149722.
 PR 20-AUG-1999; 9905-0149723.
 PR 20-AUG-1999; 9905-0149929.
 PR 23-AUG-1999; 9905-0149902.
 PR 23-AUG-1999; 9905-0149930.
 PR 25-AUG-1999; 9905-0150566.
 PR 26-AUG-1999; 9905-0150884.
 PR 27-AUG-1999; 9905-0151065.
 PR 27-AUG-1999; 9905-0151066.
 PR 27-AUG-1999; 9905-0151080.
 PR 30-AUG-1999; 9905-0151303.
 PR 31-AUG-1999; 9905-0151438.
 PR 01-SEP-1999; 9905-0151930.
 PR 07-SEP-1999; 9905-0152363.
 PR 10-SEP-1999; 9905-0153070.
 PR 13-SEP-1999; 9905-0153758.
 PR 15-SEP-1999; 9905-0154018.
 PR 16-SEP-1999; 9905-0154039.
 PR 20-SEP-1999; 9905-0154779.
 PR 22-SEP-1999; 9905-0155139.
 PR 23-SEP-1999; 9905-0155486.
 PR 24-SEP-1999; 9905-0155659.
 PR 28-SEP-1999; 9905-0156458.
 PR 29-SEP-1999; 9905-0156596.
 PR 04-OCT-1999; 9905-0157117.

PR 05-OCT-1999; 9905-0157753.
 PR 06-OCT-1999; 9905-0157665.
 PR 07-OCT-1999; 9905-0158029.
 PR 08-OCT-1999; 9905-0158232.
 PR 12-OCT-1999; 9905-0158369.
 PR 13-OCT-1999; 9905-0159293.
 PR 13-OCT-1999; 9905-0159294.
 PR 13-OCT-1999; 9905-0159295.
 PR 14-OCT-1999; 9905-0159329.
 PR 14-OCT-1999; 9905-0159330.
 PR 14-OCT-1999; 9905-0159331.
 PR 14-OCT-1999; 9905-0159637.
 PR 14-OCT-1999; 9905-0159638.
 PR 18-OCT-1999; 9905-0159584.
 PR 21-OCT-1999; 9905-0160741.
 PR 21-OCT-1999; 9905-0160767.
 PR 21-OCT-1999; 9905-0160770.
 PR 21-OCT-1999; 9905-0160776.
 PR 21-OCT-1999; 9905-0160814.
 PR 21-OCT-1999; 9905-0160815.
 PR 22-OCT-1999; 9905-0160980.
 PR 22-OCT-1999; 9905-0160981.
 PR 22-OCT-1999; 9905-0160989.
 PR 25-OCT-1999; 9905-0161404.
 PR 25-OCT-1999; 9905-0161405.
 PR 25-OCT-1999; 9905-0161406.
 PR 26-OCT-1999; 9905-0161359.
 PR 26-OCT-1999; 9905-0161360.
 PR 26-OCT-1999; 9905-0161361.
 PR 28-OCT-1999; 9905-0161920.
 PR 28-OCT-1999; 9905-0161992.
 PR 28-OCT-1999; 9905-0161993.
 PR 29-OCT-1999; 9905-0162142.

Query Match 5.38; Score 256.8; DB 21; Length 1501;
 Best Local Similarity 98.9%; Pred. No. 1,6e-49;
 Matches 269; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 3978 AAGTATGGGAATATGGAAGCTGTATGTGAG-AAGTCAATPACAGATCATTCGGG 4036
 || |||||
 Db 1230 AAATATGGGAATATGGAAGCTGTATGTGAGAAAGTCAATPACAGATCATTCGGG 1289
 |||||
 OY 4037 AATTATGATGTGACGAAGTGTGTTCTCATTTCTACCTATTACGTTAATTCGAA 4096
 |||||
 Db 1290 AATTATGATGTGACGAAGTGTGTTCTCATTTCTACCTATTACGTTAATTCGAA 1349
 |||||
 OY 4097 CGTTGGAATCATCAAAAGACCGTCCCAAAACAAATGCAATTCGATGACATTC 4156
 |||||
 Db 1350 CGTTGGAATCATCAAAAGACCGTCCCAAAACAAATGCAATTCGATGACATTC 1409
 |||||
 OY 4157 TTTGCTGATTTGATGCTATGAGTTTCAATCTAGCTACGCTATGATTTCCCTA 4216
 |||||
 Db 1410 TTTGCTGATTTGATGCTATGAGTTTCAATCTAGCTACGCTATGATTTCCCTA 1469
 |||||
 OY 4217 GATTATCAAGTTAGCCTGCGGTTTCTAAT 4248
 |||||
 Db 1470 GATTATCAAGTTAGCCTGCGGTTTCTAAT 1501
 |||||

RESULT 3
 AAC3661/c
 ID AAC3661 standard; DNA; 1194 BP.
 XX
 AC AAC3661;
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 14619.
 XX
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW Protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX

Fri Jan 17 19:40:47 2003

OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134370.
PR 16-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151067.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.

PR 15-SEP-1999; 9905-0154018.
 PR 16-SEP-1999; 9905-0154039.
 PR 20-SEP-1999; 9905-0154779.
 PR 22-SEP-1999; 9905-0155139.
 PR 23-SEP-1999; 9905-0155486.
 PR 24-SEP-1999; 9905-0155659.
 PR 28-SEP-1999; 9905-0156458.
 PR 29-SEP-1999; 9905-0156596.
 PR 04-OCT-1999; 9905-0157117.
 PR 05-OCT-1999; 9905-0157753.
 PR 06-OCT-1999; 9905-0157865.
 PR 07-OCT-1999; 9905-0158029.
 PR 08-OCT-1999; 9905-0158232.
 PR 12-OCT-1999; 9905-0158369.
 PR 13-OCT-1999; 9905-0159293.
 PR 13-OCT-1999; 9905-0159294.
 PR 13-OCT-1999; 9905-0159295.
 PR 14-OCT-1999; 9905-0159329.
 PR 14-OCT-1999; 9905-0159330.
 PR 14-OCT-1999; 9905-0159331.
 PR 14-OCT-1999; 9905-0159637.
 PR 14-OCT-1999; 9905-0159638.
 PR 18-OCT-1999; 9905-0159584.
 PR 21-OCT-1999; 9905-0160741.
 PR 21-OCT-1999; 9905-0160767.
 PR 21-OCT-1999; 9905-0160768.
 PR 21-OCT-1999; 9905-0160770.
 PR 21-OCT-1999; 9905-0160814.
 PR 21-OCT-1999; 9905-0160815.
 PR 22-OCT-1999; 9905-0160980.
 PR 22-OCT-1999; 9905-0160981.
 PR 22-OCT-1999; 9905-0160989.
 PR 25-OCT-1999; 9905-0161404.
 PR 25-OCT-1999; 9905-0161405.
 PR 25-OCT-1999; 9905-0161406.
 PR 26-OCT-1999; 9905-0161359.
 PR 26-OCT-1999; 9905-0161360.
 PR 26-OCT-1999; 9905-0161361.
 PR 28-OCT-1999; 9905-0161920.
 PR 28-OCT-1999; 9905-0161992.
 PR 28-OCT-1999; 9905-0161993.
 PR 29-OCT-1999; 9905-0162142.

Query Match 4.9%; Score 237.4; DB 21; Length 1194;
 Best Local Similarity 78.6%; Pred. No. 4.6e-45;
 Matches 353; Conservative 0; Mismatches 1; Indels 95; Gaps 2;

QY 1 AACTGTCTCTCTCCACACACAGATCTCTCTTCCGGCTGCTAGATTGGTTAAGG 60
 DB 354 AACTGTCTCTCTCCACACACAGATCTCTCTTCCGGCTGCTAGATTGGTTAAGG 296
 QY 61 ATGTGATGTTGCGAATAAATCTGTTCTTTATCATCATCATCATCATCATCATCATCAT 120
 DB 295 ATGTGATGTTGCGAATAAATCTGTTCTTTATCATCATCATCATCATCATCATCATCAT 236
 QY 121 ATGCAATCTCTTAACATACTAATAATTTCTGAAATGAAATGAAATGAAATGAAATG 180
 DB 235 ATGCAATCTCTTAACATACTAATAATTTCTGAAATGAAATGAAATGAAATGAAATG 191
 QY 181 GAATTCACACACAGATCTGAAATCTAGATTGCTGCGGAGCTACCAAAACATT 240
 DB 190 ----- 191
 QY 241 AGAAAAATGATCTTACACGATTCACCAATTTCTGGAGATTTTCTTCCAAAGCAG 300
 DB 190 -----ACGAATTCACCAATTTCTGGAGATTTTCTTCCAAAGCAG 150
 QY 301 CCATTGCTTCGAGTAGAAGAAATGTTAACTATGACTGGAGATCAGAGTAGAAGAA 360
 DB 149 CCATTGCTTCGAGTAGAAGAAATGTTAACTATGACTGGAGATCAGAGTAGAAGAA 90
 QY 361 TTGATTTTGTATGCTCATCATCTATGTTGTAATACAGATGACGAAGATCTT 420
 DB -----

DB 89 TTGATTTTGTATGCTCATCTATGTTGTAATACAGATGACGAAGATCTT 30
 QY 421 TTGAGGTTGAGAGAGAAAGCTAAGAA 449
 DB 29 TTGAGGTTGAGAGAGAAAGCTAAGAA 1

RESULT 4
 AAT39358
 ID AAT39358 standard; cDNA; 1496 BP.

AC AAT39358;

DT 09-APR-1997 (first entry)

DE Arabidopsis thaliana delta-5,7-sterol, delta7-reductase cDNA.

KW Delta-7Red; 7-dehydrocholesterol reductase; C7-unsaturated sterol;

KW pregnenolone; plant; delta-5,7 sterol, delta-7 reductase;

KW nystatin resistance; ss.

OS Arabidopsis thaliana.

FT Key Location/Qualifiers

FT CDS 76..1368

FT EP727489-AL.

PD 21-AUG-1996.

PF 14-FEB-1996; 96EP-0400301.

PR 01-JUN-1995; 95FR-0006517.

PR 15-FEB-1995; 95FR-0001723.

PA (ROUS) ROUSSEL-UCIAF.

PI Chenivresse X, Duport C, Lecain E, Pompon D;

DR WPI: 1996-372876/38.

DR P-PSDB: AAW03567.

PT Nucleic acid encoding delta-5,7 sterol delta-7 reductase - esp. of

PI A.thaliana, for producing pregnenolone-synthesizing yeast.

XX Claim 3; Page 43-46; 82pp; French.

XX A nystatin-resistant clone containing cDNA which encodes the

CC delta-5,7 sterol, delta-7 reductase enzyme of Arabidopsis thaliana

CC was isolated from a cDNA library in yeast strain Fy1679. The major

CC sterol of the untransformed yeast strain is ergosterol. RP-HPLC

CC analysis showed that in one clone the ergosterol was replaced by

CC two major sterols which did not absorb at 285 nm. The insert from

CC this clone was subcloned to localise the enzyme coding region and

CC the cDNA sequence was determined (i.e. the present sequence). The

CC enzyme, designated delta-7Red, reduces C7-unsaturated sterols. In

CC particular, it is useful in the production of pregnenolone.

XX Sequence 1496 BP; 399 A; 297 C; 340 G; 460 T; 0 other;

Query Match 4.6%; Score 224; DB 17; Length 1496;

Best Local Similarity 100.0%; Pred. No. 6.5e-42; Indels 0; Gaps 0;

Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1049 ATGATACAAATGTTCTATCAGATGTTCTTACTCAGACCTTTGGCTTTTGGGA 1108
 DB 162 ATGATACAAATGTTCTATCAGATGTTCTTACTCAGACCTTTGGCTTTTGGGA 221
 QY 1109 GAATGAGTTCAAGGACTTATCAACATATGCGCAAGACCACTTGTATGCTTGGAAA 1168
 DB 222 GAATGAGTTCAAGGACTTATCAACATATGCGCAAGACCACTTGTATGCTTGGAAA 281

Best Local Similarity 85.28; Pred. No. 1.6e-10;

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 28-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155186.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.6%; Score 79.2; DB 21; Length 629;
 Best Local Similarity 62.3%; Pred. No. 2e-08;
 Matches 142; Conservative 0; Mismatches 83; Indels 3; Gaps 1;
 QY 1050 TGTACCAATGTTTCATCAGATGTTCTGTACTCAGACCTTGGCTTTTGGAG 1109

PA (RPRG-) RPR GENCELL ASIA PACIFIC INC.

PI Hamada H;

WPI; 1999-243728/20.

PT New apoptosis-resistant virus-sensitive cell

PS Example 1; Page 38-41; 51pp; English.

The present invention describes an apoptosis-resistant virus-sensitive cell line into which an apoptosis resistance gene has been introduced. The recombinant viruses generated are capable of expressing apoptosis-associated genes. These can then be used in a variety of diseases for which the induction of apoptosis by gene transfer, or where the inhibition of harmful apoptosis, is therapeutic. The recombinant viruses are useful as vectors for gene therapy which can be applied to cancer therapy for destroying cancer cells selectively, the treatment of autoimmune diseases and graft rejection reaction, and apoptosis induction therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis-associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant 293 cell line (having an apoptosis resistant gene introduced) is established and overcomes the problem. The present sequence represents the base sequence of the plasmid pRc-1/bsr, which contains the cowpox virus bsr gene, and is used in an example from the present invention.

Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 other;

Query Match	Score	DB	Length
1.38;	63.8;	20;	6644;

Matches 275; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

[illegible]

Dy 1919 TTTCCATTTTAGTAGACCGCTAAACGGAAATGTTCTTAGCAGTATTATCTTTTATGTG 1978
||| ||| | | | | | | | | | | |
Db 3800 TT 3741

QY	1979	ATTAGTACAGAAAAC	TGTTGGT	2005
Db	3740	TTGCACTGCAAGAGG	TTATTGGAT	3714

RESULT 13
AAx33182/c
ID AAx33182 standard; DNA; 7372 BP.

AC AAX33182;

DT 25-JUN-1999 (first entry)

DE Base sequence of the plasmid pRX-Bcl-x1-bsr.

KM Compo virus; bsr: viral vector; expression; apoptosis; resistance
KM crma; bcl-2; bcl-xL; FLIP; survivin; IAP; ILP; adenovirus; cancer;
KM autoimmune disease; graft rejection reaction; inflammation;
KM inflammatory disease; ss.

05 Synthetic.

XX

2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526

XX 07 SEP 1000 0010 TD04010

XX	09-SEP-1997	07TD-03E0000E
DB		

XX
XX (PBBG-) POP GENETIC ACTA PBO

XX
PT Hamada H.

XX
DB WBT, 1000-243720/20

	Nov	9	600	+0	27	2	4
XX							
DTF							

XX Example 2; Page 41-45; 51pp; English.

The present invention describes an apoptosis-resistant virus-sensitive cell line into which an apoptosis resistance gene has been introduced. The recombinant viruses generated are capable of expressing apoptosis-associated genes. These can then be used in a variety of diseases for which the induction of apoptosis by gene transfer, or where the inhibition of harmful apoptosis, is therapeutic. The recombinant viruses are useful as vectors for gene therapy which can be applied to cancer therapy for destroying cancer cells selectively, the treatment of autoimmune diseases and graft rejection reaction, and apoptosis induction therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis-associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant 293 cell line (having an apoptosis resistant gene introduced) is established and overcomes the problem. The present sequence represents the base sequence of the plasmid pRex-Bcl-xl-bst, which contains the human Bcl-xl gene, and is used in an example from the present invention.

Sequence 7372 BP; 2353 A; 1749 C; 1649 G; 1621 T; 0 other,

Query Match	1.38;	Score 63.8;	DB 20;	Length 7372;
Best Local Similarity	43.98;	Pred. No. 0.00018;		
Matches 275; Conservative	0;	Mismatches 352;	Indels 0;	Gaps 0;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 14:13:44 ; Search time 104 Seconds
(without alignments)
943.062 Million cell updates/sec

Title: US-09-817-774-31

Perfect score: 2594
Sequence: 1 MAETVHSPIVTASMLSLA.....RAKTKQIDAIDILLICML 476

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	704.5	27.2	471	11	092228
2	553.5	21.3	620	11	008984
3	539.5	20.8	626	11	08VDM0
4	537.5	20.7	620	13	09W708
5	518	20.0	418	6	08WV1
6	495	19.1	358	11	09CXP6
7	486.5	18.8	318	11	091227
8	435.5	16.8	249	11	091Y55
9	418.5	16.1	697	5	09XVF2
10	408	15.7	292	5	09GNP2
11	399	15.4	476	3	001461
12	397.5	15.3	399	11	091WJ6
13	351.5	13.6	441	5	09N751
14	350.5	13.5	622	3	09HEL3
15	272	10.5	67	10	0945E2
16	173.5	6.7	716	5	09W2D2
					09W2D2 drosophila

17	121	4.7	460	10	09FVN0	09FVN0 lycopersico
18	119.5	4.6	284	10	09P865	09P865 oryza sativ
19	115.5	4.5	444	16	08XKR3	08XKR3 salmonella
20	112	4.3	791	10	08VXB2	08VXB2 oryza sativ
21	110.5	4.3	517	3	09HGM6	09HGM6 schizosacch
22	108.5	4.2	351	5	09U316	09U316 caenorhabdi
23	107.5	4.1	485	8	09G870	09G870 malawimonas
24	107	4.1	517	8	09B977	09B977 ceratocolen
25	106.5	4.1	524	8	09B972	09B972 ceratocolen
26	106.5	4.1	528	8	09B2800	09B2800 pyraliaella 1
27	106.5	4.1	603	17	08TL61	08TL61 methanosarc
28	105.5	4.1	525	8	063228	063228 populus tre
29	105	4.0	499	16	08YMR6	08YMR6 arabidopsis
30	105	4.0	504	10	09SVT8	09SVT8 tetrahymena
31	105	4.0	750	8	095020	095020 homo sapien
32	104.5	4.0	423	4	09H355	09H355 caenorhabdi
33	104.5	4.0	1027	5	0950C7	0950C7 limulus pol
34	104	4.0	571	8	09MLQ1	09MLQ1 buxus sp. g
35	103.5	4.0	471	8	09MH10	09MH10 lotus japon
36	103.5	4.0	569	8	09ARE8	09ARE8 lotus japon
37	103	4.0	471	5	09VT04	09VT04 drosophila
38	103	4.0	490	16	08XXR4	08XXR4 ralsstonia s
39	103	4.0	500	5	09VT03	09VT03 drosophila
40	102.5	4.0	76	3	09Y8E6	09Y8E6 neurospora
41	102.5	4.0	468	8	047441	047441 glinkgo billo
42	102.5	4.0	501	10	09FV14	09FV14 lotus japon
43	102.5	4.0	502	10	09FSH3	09FSH3 lotus japon
44	102	3.9	698	8	09SAR4	09SAR4 acrocomia a
45	101.5	3.9	462	2	051663	051663 paracoccus

ALIGNMENTS

RESULT 1

Q92228 PRELIMINARY; PRT; 471 AA.

AC Q92228; 092228; 092228

DT 01-MAY-1999 (TREMREL. 10, Created)

DT 01-MAY-1999 (TREMREL. 10, Last sequence update)

DT 01-DEC-2001 (TREMREL. 19, Last annotation update)

DE 7-DEHYDROCHOLESTEROL reductase (EC 1.3.1.21).

GN RD7R OR DHCR7.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxId=10116;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;

RA Nishino H., Ishibashi T.;

RT "Transmembrane configuration of sterol delta 7-reductase as a potential sterol sensing protein."

RL Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=SPRAGUE-DAWLEY;

RA MEDLINE=99262608; PubMed=10329655;

RX Bae S.-H., Lee J.-N., Fletzky B., Seong J.-K., Paik Y.-K.;

RT "Cholesterol biosynthesis from lanosterol. Molecular cloning, tissue distribution, expression, chromosomal localization, and regulation of rat 7-dehydrocholesterol reductase, a Smith-Lemli-Opitz syndrome-related protein."

RL J. Biol. Chem. 274:14624-14631(1999).

RN [3]

RP SEQUENCE FROM N.A.

RA Lee J.-N., Bae S.-H., Paik Y.-K.;

RT "Isolation and characterization of the rat 7-dehydrocholesterol reductase gene."

RL Submitted (Jun-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB018800; BAA34306.1; -

DR EMBL: AF071500; AAD31383.1; -

DR EMBL: AF279892; AAK69490.1; -

DR InterPro: IPR001171; ERG4_ERG24.
DR Pfam: PF01222; ERG4_ERG24; 1.
KW Oxidoreductase.
SQ SEQUENCE 471 AA; 5415 MW; BBF0CBCA4F4222EDB CRC64;

Query Match	27.28;	Score 704.5;	DB 11;	Length 471;
Best Local Similarity	36.28;	Pred. No. 5.9e-56;		
Matches 155;	Conservative 85;	Mismatches 159;	Indels 29;	Gaps 9;

Qy	14	SMLSLAFPCPEFVLLWVWVHODGSVTQTFEGF-WENGVOGLINTW---PREPLIAMKI	69
Db	38	SVIFLLLEFAR--FIVYFFIMACDQYSCSLAPILIDVATGRASLADIWAKTPRYTAAAOQL	95
Qy	70	IFCYGAEBALIOQLLP-----GKRYEGFISAGNRPVYKANGLAIVYT-----	113
Db	96	YALWSEFQVLLYSMLDPFCHRFELRGVYGVQOEGAITPAGIVNKXEYVQLOAMLITLWME	155
Qy	114	LATVIGLWVFCIFENPAIVYDR-LGEIFSAIIFGSEFICVLLIXIKGHVAPSS--DSGSCGN	171
Db	156	VNAVILSR---ESTIIIEDWVIFLLMCANILG-YAVSTRAIMIKGYLPPIISAEDCKFTGN	210
Qy	172	LIIDFYWGMELYPRIKGSFSDIKVPTNCRFGMSAVLAIVYCIKQYENKRVSDSMLVNT	231
Db	211	FIFYNVMGIEENPRIKGFEDKLEFNGRPGIVAMTILNLSPAKQOQLYGHVNTSMILVN	270
Qy	232	ILMLVYVYKFEFWMERAGVYNTDIAIDHDKRGFTICSGCLVWPVSUYTSIGMILVNHPELGT	291
Db	271	VLOAIIYVLFDEFWMETWYLUKTIIDICHNHWGLVIGGDCVWMLRYUTLTOGLVLYNHPVOLT	350
Qy	292	QLATIIIVAGSLICYINDDCDROEROEPRRTGSKCLVWGRAPSKIVASVYTTSGEFTKSL	351
Db	331	PNALGVLILLGLVGYIIFPMTHNQDRLPRRDGHCILLGKKPKALIECVTSADGCKHNHSL	390
Qy	352	LTSWGVLARHFHYVPEILSAFETVPALDENFLAIVYVIFLTLLEFDRAKRDDDCRSK	411
Db	391	LVSFGWGAHRHFNYTGDMLSLAYCLACGSGHLLPRYFYIIMTLLIHLNRDLREHRCANK	450
Qy	412	YGRYWKVLY 419	
Db	451	YGRDMERY 458	

RESULT 2	ID	008984	PRELIMINARY;	PRT;	620 AA.
AC	008984;				
DT	01-JUL-1997	(TrEMBLrel. 04, Created)			
DT	01-JUL-1997	(TrEMBLrel. 04, Last sequence update)			
DT	01-DEC-2001	(TrEMBLrel. 19, Last annotation update)			
DE	Rat NBp60.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=FISHER 344; TISSUE=LIVER.				
RA	Kawahire S., Takeuchi M., Gohshi T., Sasagawa S., Shimada M.,				
RA	Takahashi M., Abe T., Ueda T., Kuwano R., Hikawa A.;				
RT	"cDNA cloning of nuclear localization signal binding protein NBp60, a				
RT	Rat homologue of lamin B receptor, and identification of binding sites				
RT	of Human lamin B receptor for nuclear localization signals and				
RT	chromatin.";				
RL	J. Biochem. 0:0-0(0).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=FISHER 344; TISSUE=LIVER.				
RX	MEDLINE=9735934; Pubmed=9192729;				
RA	Kawahire S., Takeuchi M., Gohshi T., Sasagawa S., Shimada M.,				
RA	Takahashi M., Abe T.K., Ueda T., Kuwano R., Hikawa A., Ichimura T.,				
RA	Omata S., Horigome T.;				
RT	"cDNA cloning of nuclear localization signal binding protein NBp60, a				
RT	rat homologue of lamin B receptor, and identification of binding sites				

RT of human lamin B receptor for nuclear localization signals and chromatin.".

RL J. Biochem. 121:881-889(1997).

DR EMBL; AB002466; BAA20471.1; -.

DR InterPro: IPR001171; ERG4_ERG24.

DR InterPro: IPR002999; Tudor.

DR Pfam; PF01222; ERG4_ERG24; 1.

DR SMART; SM00333; Tudor; 1.

DR PROSITE; PS01017; ERG4_ERG24_1; 1.

DR PROSITE; PS01018; ERG4_ERG24_2; 1.

SQ SEQUENCE 620 AA; 70724 MW; 04BD9FD0C61ED5D CR664;

Query Match	21.38;	Score 553.5;	DB 11;	Length 620;
Best Local Similarity	33.08;	Pred. No. 4.8e-42;		
Matches 138;	Conservative 65;	Mismatches 176;	Indels 39;	Gaps 11;

```

QY 16 !SLIACPEFVLLMYTWVHDDGSVTO-----TFGEFMEINGVOGLINIMRPILAMKI 69
Db 221 LLIMGLPACVFLLLLLGCAQKDPGLLOFPFPLPALNELMEARVCAYILLM----- 270
QY 70 IFCYGAFEAIIQLLPGRKVEGPLSPAGNRPVYKANGLAIFVTLATYVLGLMFGIENPA 129
Db 271 FF----LQALFSLFLPVGKVEYVEGPLVDGRRLKLRNLGALAFILITLSAA-VGTAAPWDELY 325
QY 130 IYVDHIGELFSLIIGSFIFCVLLIXIGHVAPSSDS-QSCNLLIIDPYMGMEIYPRIG 188
Db 326 YLTHFLQPLAALAIYFSAVLSVYIYASLSKVPDELSPASSGNAAYDFPIGRLELPRIG- 384
QY 189 SPDIKFFTCRCGMMASMAVLAVFYCIKQYEINCKVSDSM---LVMTIMLIVYVFKFME 245
Db 385 APLKRFCELRGLLIGWVYINIMVMLAKMKVOERSAPSLAMTLVYS-FQLLYVDALME 443
QY 246 AGYWMNTIAHDRGSEYICWGCLVWPBSVYTSFGMYLVNHPVELGTQAIYIIVAGLICI 305
Db 444 EALLTMDIITHDGFEMLAGDLVWVEPFTSLQAFILVNHDDLSWPLT-SVIALKLCG 502
QY 306 YINYDC-DIQOREFPRTRNGKCLVNGRAPSKIVASYTTTSGETKTSLLTSGMWGLARHF 364
Db 503 YVIFRCANQSKNAFRKN-----PDPKLAHKITPTSTWKSLLVSGWGWGVRRPN 552
QY 365 YVBEILSAFPMVVPALFQNLPLAFYVIFITLLIFEDRAKRDDBCRSKYKWKLYACE 422
Db 553 YLSDGLMALAMSLPCGFENMLPIFYVITYITALLIHREARDHOCRRKRGGLAMEYCCQR 610

```

RESULT 3	Q8VDM0	Q8VDM0	PRELIMINARY;	PRT;	626 AA.
AC	Q8VDM0;				
DT	01-MAR-2002 (TREMBLrel, 20, Created)				
DT	01-MAR-2002 (TREMBLrel, 20, last sequence update)				
DT	01-JUN-2002 (TREMBLrel, 21, last annotation update)				
DE	Similar to Iamln B receptor.				
GN	A1505894.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eute				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mu				
OX	NCBI_TaxID=10090;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Strausberg R.				
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases				
DR	EMBL; BC021516; AAH21516.1; -				
DR	MGI; MGI:2138281; A1505894.				
DR	InterPro; IPR001171; ERG4_ERG24.				
DR	InterPro; IPR002999; Tudor.				
DR	Pfam; PF01222; ERG4_ERG24.1.				
DR	SMART; SM00333; TUDOR. 1.				
DR	PROSITE; PS01017; ERG4_ERG24.1; UNKNOWN.1.				
DR	PROSITE; PS01018; ERG4_ERG24.2; UNKNOWN.1.				
DR	Receptor.				
Q8	SEQUENCE	626 AA;	71410 MW;	BACAEE521B5F5F4DA CRC64;	

Query Match 20.8%; Score 539.5; DB 11; Length 626;
 Best Local Similarity 32.9%; Pred. No. 9.3e-41;
 Matches 135; Conservative 65; Mismatches 161; Indels 49; Gaps 12;

QY 18 LIAFPPEVILLMTMWDGDSVOTFCFEMENYQGLINIPRPTLIAMKIIFCYGAF 77
 DB 251 LHFPPPLPAL-----HE-----LMEPRVCGYLLM-----FF-----VQ 280

QY 78 AILQLLGKKEVEGEPISPAKNRPYKANGLAAYFVTLATYILGLMMFGIFNPAIYDHGE 137
 DB 281 ALFHLLPVGKAEGRPLVDGRRLQYRLNGLAFILITSA-LGAAYFVGLCYLYTHQLQ 339

QY 138 IFSALIFGSEFVCLLYTKGVADSSDS-GSCGNLIIDFYWGMELPRICKSPDIKVT 196
 DB 340 LALATGFSVLLSALVLYRSLRPRELSPASSGNAVYDFIGRELNPRIQ-ANDLAFPC 398

QY 197 NCRGMSMAVLAATYCCIKOYEINGKSDSM---LVNTIIMLVYVTFEWMAGIYNTMD 253
 DB 399 ELRPGELIGWVYINLMLMEKKIQERRAPSLAMILVNS-FQLLYVDALMNEEALLTSM 457

QY 254 IAHDRGFYICWGLVWVPSYVTSFGMYLVNHPVELGTOLAIYILVAGILCIYINYDC-D 312
 DB 458 IMHGFEGMLAFGDLVWVPTYSLOAFYLVSHPHDLSPMLA-SYIILKLCGYIIFRCAN 516

QY 313 RQROBFRTNGKCLVWGRAPSKIYASVYTTTSGETKTSLLTSGWGLARHVFYETISA 372
 DB 517 SOKNAFRKN-----PTDPKLAHLKTIHTSTGKSLVSGWGVGRHNPVYGLDIMA 566

QY 373 FFWVTVPALFDMFLAYFYVIFLTLFPDRAKDDRCRSKYKYLCKE 422
 DB 567 LAMSLPCGFHLLPFIYIYFTALLIHRDEHQCRRKYGLAMEKCYCOR 616

RESULT 4
 Q9W708 PRELIMINARY; PRT: 620 AA.

AC Q9W708;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Lamin B receptor.
 OS P58 GENE.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipodae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99323925; PubMed=10393814;
 RA Gajewski A., Krohne G.;
 RT "Subcellular distribution of the Xenopus p58/Lamin B receptor in
 oocytes and eggs";
 RL J. Cell Sci. 112:2583-2596(1999).
 DR EMBL: Y17842; CABA4317.1; -;
 DR InterPro: IPR000515; BPD_transp.
 DR InterPro: IPR001171; ERG4_ERG24.
 DR InterPro: IPR002999; Tudor.
 DR Pfam: PF01222; ERG4_ERG24; 1.
 DR SMART: SM00333; TUDOR; 1.
 DR PROSITE: PS00402; BPD_TRANS_P-INN-MEMBR; UNKNOWN_1.
 KM Receptor.
 SQ SEQUENCE 620 AA; 71128 MW; B70179E01F57C5EA CRC64;

Query Match 20.7%; Score 537.5; DB 13; Length 620;
 Best Local Similarity 31.0%; Pred. No. 1.4e-40;
 Matches 135; Conservative 71; Mismatches 177; Indels 53; Gaps 9;

QY 3 EYVHSPITYASMUSLAF--CPEVILLMTMWDGDSV-----QTGFGE 49
 DB 210 KNEENPLAIGGALGFLFMWCVP--ALLYLLVVCQGYTSGPVFVELLDIVGGEF-- 265
 QY 50 NGVGLINIPRPTLIAMKIIFCYGAFEALILQLLPGRKVEGPISPAGNRPYKANGLA 109

DB 266 -----VIMTL-----QVLLYLLPGLKVVDDGQOLKNRKLRIISGCSA 304

QY 110 YEVTLATYILGLMMFGIFNPAIYDHLGEIFLSALIFGSEFVCLLYTKGVADSS--SDSG 167
 DB 305 FFLTAIIMAGMKRYEINFLYIEFHYLOFASATLSFSLISLYLVRSKYVPRELSMA 364

QY 168 SCGNLIIDFYWGMELPRICKSPDIKVTNCRGMSMAVLAATYCIKOEIN-GKVD 225
 DB 365 NSGNFIYKFWGRIENRIG-NLDLVFVYVIRQALMSWVILINIMLFAEKYHKMDPSL 423

QY 226 SMLVNTIIMLVYVTFEWMAGIYNTMDIAHDGGEFYICWGLVWVPSYVTSFGMYLVN 285
 DB 424 SMLVNSFOLLVYLDGFWNNEFYFLMSPDIYRDOFGFLAFLAFAFTYSLOTIYLVNN 483

QY 286 PVELGTOLAIYILVAGILCIYINYDCRQROBFRTNGKCLVWGRAPSKIYASVYTTTSGE 345
 DB 484 PVDLSRQASAIYALFELGYIIRGANNOKCAFRON-----PDDPRLSHLKIPT 533

QY 346 TKTSLITSGWGLARHVFYETISAFTVYVATLNDNLAIFYVIFLTLFPDRAKRD 405
 DB 534 SAGSKLITSGWGVGRHNPVYGLDILMALAMCLACGPHLLPFYVIFLTLIDRAARDE 593

QY 406 DRCRSKYKYLCKE 421
 DB 594 QRCREKYGDMDKYCC 609

RESULT 5
 Q8WV1 PRELIMINARY; PRT: 418 AA.

AC Q8WV1;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE C-14 sterol reductase.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX PubMed=11784322;
 RA Roberti R., Bennati A.M., Galli G., Caruso D., Maras B., Aisa C.,
 RT "Cloning and expression of sterol delta14-reductase from bovine
 liver";
 RL Eur. J. Biochem. 269:283-290(2002).
 DR EMBL: AF039681; AAK91505.1; -;
 DR InterPro: IPR001171; ERG4_ERG24.
 DR Pfam: PF01222; ERG4_ERG24; 1.
 DR PROSITE: PS01017; ERG4_ERG24_1; UNKNOWN_1.
 DR PROSITE: PS01018; ERG4_ERG24_2; UNKNOWN_1.
 SQ SEQUENCE 418 AA; 46751 MW; 30DE297CFE09568C CRC64;

Query Match 20.0%; Score 518; DB 6; Length 418;
 Best Local Similarity 35.3%; Pred. No. 5.4e-39;
 Matches 129; Conservative 57; Mismatches 153; Indels 26; Gaps 10;

QY 67 WKIIFC--YGAFEALILQLLPGRKVEGPISPAGNRPYKANGLAAYFVTLATYILGLMMFG 124
 DB 61 WMLLCLLWGLQALALYLLPARKVAEGGLKDKSRIRYPTNFQALVLT-ALLVGL---G 116

QY 125 INRPAIYDHLGEIFLSALIFG----SFIFCVLLYTKGVADSS--SDSGSCGNLIIDFY 178
 DB 117 V-SAGIPLSALPEMLPLAFAATLTAFTISLLYIKALAPASALAPGNGSNLIYDFEL 175

QY 179 GMEIYPRICKSPDIKVTNCRGMSMAVLAATYCIKOEINGKSDSMVNTIIMLVY 238
 DB 176 GRELNPRI-CSPDFKFCFLRGLIGWVLIINALLIQEELGSGSLMMWMLVNGFOLLIV 234

QY 239 TKFWMAGIYNTMDIAHDGGEFYICWGLVWVPSYVTSFGMYLVNHPVELGTOLAIYIL 298

DB 235 GDLMTWEEAVLTITMDIITHGFGMLAFGLDPAWPFYISQAGFLYHPRLDPLMPPLASFIC 294
 QY 299 VAGILCIYINDCDROQROEFRTNGKCLWGRAP-S-KIVASTTTTSGEKTSLLTSGMW 357
 DB 295 LINAAGYIFRGANSOKNFRKN-----PSDPRVADELTISTATGRRLTY-SGMW 343
 QY 358 GLARHHYVEILLASFFWVPALFDFNLAFYVIFLTLLFPRAKDDDCRSKYKWK 417
 DB 344 GMRHHPYIGDLIMALMSLPCGVFHLPLPYFLYFTALLVHREDEROCROKYLAWH 403
 QY 418 LYCEK 422
 DB 404 EYCR 408

RESULT 6
 Q9CXP6 PRELIMINARY; PRT; 358 AA.

AC 09CXP6: 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE 3110041018RIK protein.
 GN 3110041018RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC HEAD;
 RX MEDLINE=21085660; PubMed=1117851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Glassi C., King B., Kochava H.,
 RA Knehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlaker C., Wilmink L.,
 RA Wysshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK014166; BAB29187.1; -;
 DR MGD: MGI:1920416; 3110041018RIK.
 DR InterPro: IPR001171; ERG4_ERG24.
 DR Pfam: PF01222; ERG4_ERG24_1; UNKNOWN_1.
 DR PROSITE: PS01017; ERG4_ERG24_1; UNKNOWN_1.
 DR PROSITE: PS01018; ERG4_ERG24_2; 1.
 SQ SEQUENCE 358 AA; 39719 MW; 0DD8D4AB2DE0275C CRC64;

Query Match 19.1%; Score 495; DB 11; Length 358;
 Best Local Similarity 34.6%; Pred. No. 5, 6e-17;
 Matches 113; Conservative 58; Mismatches 134; Indels 22; Gaps 8;

QY 102 YKANGLAAYVTATYIGLWFGIFNPAYVDHLGELISALIFG---SEIFCVLLYIKG 157
 DB 38 YVGFQALVLT-ALLMGL---GV-SVGLPLGALPGLPLPLATITLTSFISLLYAKA 92
 QY 158 HVAPSS--SSGSGGNLIDPYMGMEYPRIGSFIDKYVTNCRFGMSAAVLAIVYCIK 215
 DB 93 LVAPASALAGGNSGNSMDYDFLGRLELNPRIQ-SFDEKYPCELRPGILGWVFINALLMQ 151

QY 216 QYEINGKVSMDLVNTILMLVYTKFFWMEAGYWNMDIAHRCGFYICWGCLWVPSSVY 275
 DB 152 EAELRGSPSLAMLVNGLVNGFQLLYGDALMYEESVLTITMDIITHGFGMLAFGLDPAWPFY 211
 QY 276 TSPGMVYNHPELGTOLAIYILVAGILCIYINDCDROQROEFRTNGKCLWGRAPSKI 335
 DB 212 SLQAFLLYHPQPLGLPALLICLLKLVGYIFPGANSOKNFRKNSDPSVAG----- 265
 QY 336 VASTTTTSGEKTSLLTSGMWGLARHHYVEILLASFFWVPALFDFNLAFYVIFLT 395
 DB 266 LETIPTARGRO---LVVSGMGVNRHNPYIGDLIMALMSLPCGLSHLLPYEVLVFTA 321
 QY 396 LLEPRKRDRCRSKYKWKLYCEK 422
 DB 322 LIVREARDEQCLQKGRAMQETCKR 348

RESULT 7
 Q91227 PRELIMINARY; PRT; 318 AA.

AC 091227: 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical 36.4 kDa protein (Fragment).
 GN A1505894.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: BC010261; AAH10261.1; -;
 DR MGD: MGI:2138281; A1505894.
 DR InterPro: IPR001171; ERG4_ERG24.
 DR Pfam: PF01222; ERG4_ERG24_1.
 DR PROSITE: PS01017; ERG4_ERG24_1; UNKNOWN_1.
 DR PROSITE: PS01018; ERG4_ERG24_2; UNKNOWN_1.
 KW Hypothetical protein.
 FT NON_TER
 SQ SEQUENCE 318 AA; 36384 MW; 91A05EB39274E398 CRC64;

Query Match 18.8%; Score 486.5; DB 11; Length 318;
 Best Local Similarity 34.5%; Pred. No. 2, 9e-36;
 Matches 111; Conservative 59; Mismatches 133; Indels 19; Gaps 8;

QY 106 GLAAAYVTATYIGLWFGIFNPAYVDHLGELISALIFGSEIFCVLLYIKGHVAPSSSD 165
 DB 1 GLVAFILTSAA-LGAAYFWGVELCYLTHFLQALAAATGFSVLLSAYLVKSLRAPREEL 59
 QY 166 S-GSGGNLIDPYMGMEYPRIGSFIDKYVTNCRFGMSAAVLAIVYCIQYINRKYV 224
 DB 60 SPASGNAVYDFEIRLEINPRIG-AFDLKFCELRPGILGWVINLVLMEMKIOERAA 118
 QY 225 DSM--LVNTIIMLYVYTKFFWMEAGYWNMDIAHRCGFYICWGCLWVPSSVY 281
 DB 119 PELAMILVNS-FQLLYVDALMNEALLTSMIDIMHDGFGMLAFGLDPAWPFYISLQAFY 177
 QY 282 LVNHPELGTOLAIYILVAGILCIYINDC-DROQROEFRTNGKCLWGRAPSKIYASYT 340
 DB 178 LVSHPHDLSWPLA-SVIALKLCGVIVIFRANCSOKNFRKN-----PTDPKLAHL 226
 QY 341 TTSGETKTSLLTSGMWGLARHHYVEILLASFFWVPALFDFNLAFYVIFLTLLFDR 400
 DB 227 KTIHTSTGKSLIVSGMWGEVRHNPYIGDLIMALMSLPCGNHLLPYIIVFYALLIHR 286
 QY 401 AKRDDRCRSKYKWKLYCEK 422
 DB 287 EARDEHQCRKRYGLAMEKTCOR 308

```

RESULT 8
Q91YS5 PRELIMINARY; PRT; 249 AA.
AC Q91YS5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical 29.0 kDa protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC014835; AAI14835.1;
DR InterPro: IPR001171; ERG4_ERG24.
DR Pfam: PF01222; ERG4_ERG24; 1.
DR PROSITE: PS01018; ERG4_ERG24_2; UNKNOWN_1.
KW Hypothetical protein.
FT NON_TER
SQ SEQUENCE 249 AA; 28955 MW; 86CF29C0E75341BF CRC64;

Query Match 16.8%; Score 435.5; DB 11; Length 249;
Best Local Similarity 35.7%; Pred. No. 1e-31;
Matches 90; Conservative 50; Mismatches 95; Indels 17; Gaps 6;

QY 175 DEYWGMEIYRIGKSPDIKFTNCRFGMSMAVLAVTYCKQYEINGKVSDSM---LVNT 231
DB 1 DFFGRLNPLRG-APLKFECFELRPLIGWVYINVLMLMEKIDERRAPSLAMILVNS 59
QY 232 ILMVYTKFEEWAGYWNMDIAHNRGFGYICWGLVWVPSYTSFGVLYVHNPVLTG 291
DB 60 -FQLLYVDALWMEALALTSDIMHDSFGMLAFGLDLYWVPFTYSQAFLVSHPHDLSW 118
QY 292 QLAIIIVAGILCIYINYDC-DNRQREFRTNCKLVGRAPSKIYASVTTTSGETKTSL 350
DB 119 PLA-SVIAIKRLCGYVIFRCANSQKNAFRKN-----PDDPKLAHLKTIHTSGKS 167
QY 351 LITSGMGLARHPHYVEIISAFPMVPALEDFNPLAFVYIFLTLFLDPAKDDDCRS 410
DB 168 LTVSGMGEFVRHPNYLDDLIMALMSLPCGFNHLDPFYITFTALLIHREARDEHQCR 227
QY 411 KYGRYMKLYCEK 422
DB 228 KYGLAMEKCYCOR 239

RESULT 9
Q9XVF2 PRELIMINARY; PRT; 697 AA.
AC Q9XVF2;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE B0250.9 protein.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Lloyd C.R.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=9069613; PubMed=9851916;
RX none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).

```

```

DR EMBL: Z81453; CAB03797.1; -.
DR InterPro: IPR002106; AARNA_1lgaseII.
DR InterPro: IPR002892; DUF40.
DR InterPro: IPR001171; ERG4_ERG24.
DR Pfam: PF01838; DUF40.1.
DR Pfam: PF01222; ERG4_ERG24; 1.
DR PROSITE: PS00339; AA-TRNA LIGASE II 2; UNKNOWN_1.
SQ SEQUENCE 697 AA; 79878 MW; 4B93C5E1A5D6C8 CRC64;

Query Match 16.1%; Score 418.5; DB 5; Length 697;
Best Local Similarity 29.3%; Pred. No. 1.2e-29;
Matches 122; Conservative 76; Mismatches 200; Indels 19; Gaps 9;

QY 10 VTYASMLSLIACPPVILLMTWTHQD-----SYQTGFGEWNGVGLIMYRPILI 65
DB 286 VSAQMWALLLIIVPPATPELFYSISIHGLFVPTTALFLRP-PLVLCVPPVW---DTV 341
QY 66 AMKIFCYGAFPAIQLDILPGKRVESGISPAGRPPYKANGLAAYFVLATYGLMFGI 125
DB 342 AMKFSVHCAIQILFVWVPHDQAL-VKSSAGDQ-MREVNSFSCILICLLVGLASAGV 399
QY 126 FNDALVYDHLGEIFSAIIFGSEIFCVLLYIKGVAPSSDSCGNLIIDFYMGELYPH 185
DB 400 YRGDLVYLHFNSEII-LIPA--IFAVLLW----AALIAHYHGVTYTISEFWGIENHPK 451
QY 186 ICKSPDIKFTNCRFGMSMAVLAVTYCKQYEINGKVSDSMVLNTIIMLYVTKFEWME 245
DB 452 I-LDDIDKSFITRTFEVLPVLPVLSAMVHKKITVGGIISTSLVCLSSVQLLYIFQPHNE 510
QY 246 AGYWTMTDIAHNRGFGYICWGLVWVPSYTSFGVLYVHNPVLTGQALYIYLVAGILCI 305
DB 511 DFLNSLDSKRCDFGFRYIMADVFLGPIITYSPTVILVATNSVVISNCLCAVALISM 570
QY 306 YINYDCDNRQREFRTNCKLVGRAPSKIYASVTTTSGETKTSLLTSGMGLARHPHY 365
DB 571 VETACDRQKYEFRKSKGLKVGDAEFISAKYRTDSDGANTNLLTSGHGVCRHPNY 630
QY 366 VPEIISAFPMVPALEDFNPLAFVYIFLTLFLDPAKDDDCRSKYKMKLYCEK 422
DB 631 ASEAITFAFSAPQGPSTIAHIPSILFVLVVARAFDENRCLIKYQWMAQYCK 687

RESULT 10
Q9GNE2 PRELIMINARY; PRT; 292 AA.
AC Q9GNE2;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Putative sterol C-14 reductase (Fragment).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RA Bhavani T.P.; Kasbekar D.P.;
RT "Genomic sequence of the putative sterol C-14 reductase of
RT Dictyostelium discoideum.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Prakash A.; Prasanna B.T.; Kasbekar D.P.;
RT "SSG792 cDNA encoding the putative sterol C-14 reductase of
RT Dictyostelium discoideum.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF308471; AAG30272.1; -.
DR EMBL: AF308470; AAG30271.1; -.
DR InterPro: IPR001171; ERG4_ERG24.
DR Pfam: PF01222; ERG4_ERG24; 1.
DR PROSITE: PS01017; ERG4_ERG24_1; UNKNOWN_1.
FT NON_TER
SQ SEQUENCE 292 AA; 34205 MW; 4A7F42DA72F1C53E CRC64;

```


RESULT 14			
09HEL3			
ID	09HEL3	PRELIMINARY;	PRT; 622 AA.
AC	09HEL3;		
DT	01-MAR-2001 (TREMBLrel_16, Created)		
DT	01-DEC-2001 (TREMBLrel_19, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel_20, Last annotation update)		
DE	Probable sterol C-24 reductase.		
GN	12F11.140.		
OS	Neurospora crassa.		
OC	Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;		
OC	Sordariales; Sordariaceae; Neurospora.		
OX	NCBI_TaxID=5141;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RA	Schulte U., Aign V., Hohelsel J., Brandt P., Partmann B., Holland R.,		
RA	Yakutana G., Mewes H.W., Mannhaupt G.;		
RL	Submitted (Dec-2000) to the EMBL/GenBank/DBJ databases.		
RL	[2]		
RP	SEQUENCE FROM N.A.		
RA	German Neurospora genome project;		
RA	Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.		

Query Match	10.5%	Score 272;	DB 10;	Length 67;
Best Local Similarity	85.5%;	Pred No.	1.9e-17;	
Matches 47;	Conservative 7;	Mismatches 1;	Indels 0;	Gaps 0;
QY	368	ELLSAEFFVYPALFDNLFAYFVLTLLDPRAKRDDRCRSYGYKMYLCEK	422	
DG	3	QLLSAEFWVPALFNHFLPEFFVILTLTLDPRAKRDDRCRSYGYKMYVPEEK	57	

Fri Jan 17 19:40:59 2003

us-09-817-774-31.rspt

Job time : 108 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 14:13:05 ; Search time 13 Seconds

(Without alignments)
1518.672 Million cell updates/sec

Title: US-09-817-774-31

Perfect score: 2594

Sequence: 1 MAETVHSPIVTYASMLSLA.....RAKTKMQIDAIDILLICML 476

Scoring table: BIOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 112892 seqs, 41476328 residues

112892

Total number of hits satisfying chosen parameters:

112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt.40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2312	89.1	432	1	ST7R_ARATH
2	710.5	27.4	471	1	DHC7_MOUSE
3	698.5	26.9	475	1	DHC7_HUMAN
4	554	21.4	637	1	LBR_CHICK
5	544	21.0	615	1	LBR_HUMAN
6	503.5	19.4	430	1	ER24_ASCIM
7	500.5	19.3	424	1	ER24_SCHPO
8	489.5	19.3	418	1	ER24_HUMAN
9	473.5	18.3	490	1	ER24_NEUCR
10	470.5	18.1	369	1	ER24_ARATH
11	459.5	17.7	485	1	ER24_FUSSO
12	458.5	17.7	512	1	ER24_SEPLY
13	424.5	16.4	438	1	ER24_YEAST
14	371	14.3	473	1	ER24_YEAST
15	364.5	14.1	453	1	ST31_SCHPO
16	121	4.7	460	1	AT13_LYCES
17	120.5	4.6	524	1	COX1_BETVU
18	115.5	4.5	527	1	COX1_ARATH
19	108	4.2	599	1	THIY_YEAST
20	106	4.1	354	1	OPSD_CARAU
21	106	4.1	474	1	NU4M_PARTE
22	105	4.0	504	1	AT14_ARATH
23	104.5	4.0	1418	1	CELL_CAEEL
24	104	4.0	632	1	AFUB_LOCM1
25	104	4.0	632	1	AFUB_HAEIN
26	102	3.9	438	1	SECV_MERYA
27	101.5	3.9	514	1	AT12_LYCES
28	101.5	3.9	522	1	COX1_MARPO
29	101.5	3.9	538	1	THIY_YEAST
30	101	3.9	354	1	OPSD_POERE
31	100.5	3.9	485	1	YIDL_ECOLI
32	100	3.9	354	1	OPSD_CYPCA
33	98.5	3.8	558	1	CXIB_PARDE

34	97.5	3.8	173	1	NU6M_LATCH	003175 Latimeria C
35	96.5	3.7	524	1	COX1_ORYSA	P14578 Oryza sativ
36	96.5	3.7	528	1	COX1_MAZE	P08742 zea mays (m
37	96.5	3.7	530	1	COX1_SORBI	P05502 sorghum bic
38	96	3.7	354	1	OPSD_SALPY	P09423 salaria pav
39	96	3.7	459	1	YRPL_YEAST	P53584 saccharomyc
40	95.5	3.7	389	1	LGT_MYCPN	P75547 mycoplasma
41	95.5	3.7	524	1	COX1_WHEAT	P08741 triticum ae
42	95	3.7	514	1	AT12_ARATH	P09218 arabidopsis
43	94	3.6	687	1	AFUB_ACTPL	Q44123 actinobacil
44	94	3.6	702	1	NU5C_POAPR	Q32880 poa pratens
45	93	3.6	354	1	OPSD_ORYLA	P87369 oryzias lat

ALIGNMENTS

RESULT 1
ST7R_ARATH STANDARD: PRT: 432 AA.
ID ST7R_ARATH
AC 09LDT6: Q38930: (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 7-dehydrocholesterol reductase (EC 1.3.1.21) (7-DHC reductase) (Sterol
delta-7-reductase) (Dwaf5 protein)
GN DW5 OR ST7R OR AT1G50430 OR F11F12.21.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE-Seeding;
RX MEDLINE=96210024; PubMed=8631902;
RA Lecain E., Chenivasse X., Spagnoli R., Pompon D.;
RT "Cloning by metabolic interference in yeast and enzymatic
characterization of Arabidopsis thaliana sterol delta 7-reductase.";
RL J. Biol. Chem. 271:10866-10873(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. WS-2;
RX MEDLINE=20223137; PubMed=10758495;
RA Choe S., Tanaka A., Noguchi T., Fujioke S., Takatsuto S., Ross A.S.,
RA Tax F.E., Yoshida S., Feldmann K.A.;
RT "Lesions in the sterol delta reductase gene of Arabidopsis cause
dwarfism due to a block in brassinosteroid biosynthesis.";
RL Plant J. 21:431-443(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-CV. Columbia;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chen Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Cressy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hultzer L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetskaia I., Kutz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziani A.,
RA Miltischer J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pail G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Uteback T., Van Aken S., Vaysberg M., Vayskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
thaliana.";
RL Nature 408:816-820(2000).
-1- FUNCTION: Production of cholesterol by reduction of C7-C8 double

CC bond of 7-dehydrocholesterol (7-DHC). Lesions in the gene coding
 CC for the enzyme cause dwarfism.
 CC -1- CATALYTIC ACTIVITY: Cholesterol + NADP(+) = cholesta-5,7-dien-3-
 CC beta-ol + NADPH.
 CC -1- PATHWAY: Sterol biosynthesis. Also participate in brassinosteroid
 CC (BR) biosynthetic pathway.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
 CC reticulum.
 CC -1- SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U49398; AAC49278.1; -;
 DR EMBL: AF239701; AAF63498.1; -;
 DR EMBL: AC012561; AAF87888.1; -;
 DR InterPro: IPR001171; ERG4_ERG24.
 DR Pfam: PF01222; ERG4_ERG24; 1.
 DR PROSITE: PS01017; STEROL_REDUCT_1; 1.
 DR PROSITE: PS01018; STEROL_REDUCT_2; 1.
 KW Sterol biosynthesis; Oxidoreductase; NADP; Transmembrane;
 KM Endoplasmic reticulum.
 FT TRANSMEM 12 34 POTENTIAL.
 FT TRANSMEM 64 86 POTENTIAL.
 FT TRANSMEM 107 126 POTENTIAL.
 FT TRANSMEM 136 155 POTENTIAL.
 FT TRANSMEM 195 212 POTENTIAL.
 FT TRANSMEM 227 249 POTENTIAL.
 FT TRANSMEM 261 283 POTENTIAL.
 FT TRANSMEM 287 309 POTENTIAL.
 FT TRANSMEM 371 393 POTENTIAL.
 FT VARIANT 257 257 D -> N (IN DMF5-4; DWARF PLANT).
 FT VARIANT 117 117 Y -> H (IN REF. 1).
 FT CONFLICT 308 308 N -> K (IN REF. 1).
 FT CONFLICT 391 392 MISSING (IN REF. 1).
 SQ SEQUENCE 432 AA; 49591 MW; 5458E495BE1E4B0 CRC64;

Query Match 89.1%; Score 2312; DB 1; Length 432;
 Best Local Similarity 99.8%; Pred. No. 7.1e-163;
 Matches 421; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAETVSPITVYASMLSLAFCEPPVILMYWHDGSVYTFGFWENGVOGLINIMP 60
 DB 1 MAETVSPITVYASMLSLAFCEPPVILMYWHDGSVYTFGFWENGVOGLINIMP 60
 QY 61 RPLLIAMKIFFCYGAFFAIIQLLPKRVGSPISPAENRPPYKANGLAAYFTLATYGL 120
 DB 61 RPLLIAMKIFFCYGAFFAIIQLLPKRVGSPISPAENRPPYKANGLAAYFTLATYGL 120
 QY 121 WMFGFNPALVYDHLGEIFESALIFGSIFCVLLYKINGHVAPESSDSSGSGNLIIDPYWGM 180
 DB 121 WMFGFNPALVYDHLGEIFESALIFGSIFCVLLYKINGHVAPESSDSSGSGNLIIDPYWGM 180
 QY 181 ELXPRIKGSDFIVFNCRGMMKSAVLAITYCIKOYEINGKSDSMVNTIMLYYVRK 240
 DB 181 ELXPRIKGSDFIVFNCRGMMKSAVLAITYCIKOYEINGKSDSMVNTIMLYYVRK 240
 QY 241 ELPRIKGSDFIVFNCRGMMKSAVLAITYCIKOYEINGKSDSMVNTIMLYYVRK 240
 DB 241 ELPRIKGSDFIVFNCRGMMKSAVLAITYCIKOYEINGKSDSMVNTIMLYYVRK 240
 QY 241 FPMWEGVYNTMDIAHDSRGFYICMGCLVWPVSVYSPGYLVNHNVEIGTOLAIYVLA 300
 DB 241 FPMWEGVYNTMDIAHDSRGFYICMGCLVWPVSVYSPGYLVNHNVEIGTOLAIYVLA 300
 QY 301 GILCIYINDCDROREFRRTNGKCLWGRAPSKIIVASYTTTSGETKTSLLTSGMWGLA 360
 DB 301 GILCIYINDCDROREFRRTNGKCLWGRAPSKIIVASYTTTSGETKTSLLTSGMWGLA 360
 QY 361 RHFHYVEIISAFFWYVPALENFVLAIFYITLLLPDRAKDDDRCSKYGKTKLYC 420
 DB 361 RHFHYVEIISAFFWYVPALENFVLAIFYITLLLPDRAKDDDRCSKYGKTKLYC 420

QY 421 EK 422
 DB 421 EK 422

RESULT 2
 DHC7_MOUSE
 ID DHC7_MOUSE STANDARD; PRT; 471 AA.
 AC 088455;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 7-dehydrocholesterol reductase (EC 1.3.1.21) (7-DHC reductase) (Sterol
 DE delta-7-reductase).
 GN DHC7.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=98318632; Pubmed=9653161;
 RA Fitzky B.U., Witsch-Baumgartner M., Erdel M., Lee J.N., Paik Y.K.,
 RA Glossmann H., Utermann G., Moebius F.F.;
 RA "Mutations in the delta7-sterol reductase gene in patients with the
 RA Smith-Lemli-Opitz syndrome.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8181-8186(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Production of cholesterol by reduction of C7-C8 double
 CC bond of 7-dehydrocholesterol (7-DHC).
 CC -1- CATALYTIC ACTIVITY: Cholesterol + NADP(+) = cholesta-5,7-dien-3-
 CC beta-ol + NADPH.
 CC -1- PATHWAY: Sterol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
 CC reticulum.
 CC -1- SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF057368; AAC40164.1; -;
 DR EMBL: BC006854; AAH06854.1; -;
 DR MGD: MGI:1298378; DHC7.
 DR InterPro: IPR001171; ERG4_ERG24.
 DR Pfam: PF01222; ERG4_ERG24; 1.
 DR PROSITE: PS01017; STEROL_REDUCT_1; 1.
 DR PROSITE: PS01018; STEROL_REDUCT_2; 1.
 KW Sterol biosynthesis; Cholesterol biosynthesis; Oxidoreductase; NADP;
 KM Transmembrane; Endoplasmic reticulum.
 FT TRANSMEM 33 55 POTENTIAL.
 FT TRANSMEM 147 169 POTENTIAL.
 FT TRANSMEM 174 196 POTENTIAL.
 FT TRANSMEM 298 320 POTENTIAL.
 FT TRANSMEM 325 347 POTENTIAL.
 FT TRANSMEM 410 432 POTENTIAL.
 SQ SEQUENCE 471 AA; 53918 MW; 6B1BC356CC39290 CRC64;

Query Match 27.4%; Score 710.5; DB 1; Length 471;
 Best Local Similarity 36.6%; Pred. No. 3.4e-45;
 Matches 157; Conservative 85; Mismatches 158; Indels 29; Gaps 9;

QY 13 ASMSLAFCEPPVILMYWHDGSVYTFGFWENGVOGLINIMP 60
 DB 13 ASMSLAFCEPPVILMYWHDGSVYTFGFWENGVOGLINIMP 60

Db 37 ASIIILLPAP--FTVYFIMACDQYSCSLAPALDIATGASHLADIWAKTPPYTANAAQ 94
 QY IIFCYGAFEAIIOLLPP-----GKREGEPISPAGNRPYKANGLAAYVT---- 113
 Db 95 LYALMVSQVILLYSWLPDPCHEFLPBGYGVQGEAIFPAGVNVKNVNGLOAMITHTLM 154
 QY 114 -LATYGLMWFGEINPAIVYDH-IGEIFSALIFSSIFCVLLYTKGHVAPSS--DSSGSC 170
 Db 155 FVNATILSM---FSPFIIFDNWIPILMVCANILG-YAVSTFAMIKGYLFPSADCKFTG 209
 QY 171 NLIIDEFWGMELYPRIKGSFPIKFTNCRFGMSMAVLAIVYTCIKOIEINKVSDSMLVN 230
 Db 210 NFFYVMGIEFNPRIKGFPEFKLEFNCRPISIVAMTLLINSEFAKQOELGCHVNSMLV 269
 QY 221 TILMLVYTKFEFWMEAGVWNTMDIAHDRGFFYICMGCLVWPVSYYTSPGMVHPVGLG 290
 Db 270 NVLQAIIVLDFEWNMTWLKIDICHDFWYLGNGDCVWLPYXTLOGLIVLHPVQLS 329
 QY 291 TOLAIYIIIVAGILCIYINWYCDROQERRTNGKCLVWGRAPSKIVASYTTSGEKTSL 350
 Db 330 TPNALGILLGLVGYIIFRMTNHOKDLFRTRDGRCLIMGKPKALIECSYTSADGLKHSK 389
 QY 351 ILTSGMGLARHFVHPVPEILSAPFVTVAPALDNFLAFYVYVFLLLFDRAKRDRCRS 410
 Db 390 LTVSGFWGVARHFVNTGDLMSLAVCLAAGGHLIPFYIITMTILLTHRLRDEHRCAN 449
 QY 411 KYGKWKLY 419
 Db 450 KYGRMERY 458
 RESULT 3
 DHCT_HUMAN STANDARD: PRT: 475 AA.
 ID DHCT_HUMAN
 AC Q9UBW7: 060492: 060717:
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 7-dehydrocholesterol reductase (EC 1.3.1.21) (7-DHC reductase) (Sterol
 delta-7-reductase) (Putative sterol reductase SR-2).
 GN DHCR7 OR D7SR.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANTS SLOS LEU-119; ARG-244 AND CYS-248.
 RX PubMed=9683613;
 RA Waterham H.R., Wijnburg F.A., Hennekam R.C.M., Vreken P.,
 RA Poll-The B.T., Dordani L., Duran M., Jila P.E., Smeitink J.A.M.,
 RA Wevers R.A., Wanders R.J.A.;
 RT "Smith-Lemli-Opitz syndrome is caused by mutations in the
 RT 7-dehydrocholesterol reductase gene.";
 RL Am. J. Hum. Genet. 63:329-338(1998).
 RN [2]
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RX TISSUE=Liver;
 RC MEDLINE=98132689; PubMed=9465114;
 RA Moebius F.F., Fitzky B.U., Lee J.N., Paik Y.K., Glossmann H.,
 RT "Molecular cloning and expression of the human delta7-sterol
 RT reductase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:1899-1902(1998).
 RN [3]
 RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
 RX MEDLINE=99097347; PubMed=9878250;
 RA Holmer L., Pezhaman A., Norman H.J.;
 RT "The human lamin B receptor/sterol reductase multigene family.";
 RL Genomics 54:469-476(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

RN [5]
 RP SEQUENCE OF 14-475 FROM N.A.
 RC TISSUE=Liver;
 RX PubMed=9634533;
 RA Wassif C.A., Maslen C., Kachilele-Linjeweile S., Lin D., Linck L.M.,
 RA Conner W.E., Steiner R.D., Porter F.D.;
 RT "Mutations in the human sterol delta 7-reductase gene at 11q12-13
 RT cause Smith-Lemli-Opitz syndrome.";
 RL Am. J. Hum. Genet. 63:55-62(1998).
 RN [6]
 RP VARIANTS SLOS S-51; M-93; P-99; P-157; V-247; I-326; W-352; S-380;
 RP C-404 AND S-410.
 RX MEDLINE=98318632; PubMed=9653161;
 RA Fitzky B.U., Witsch-Baumgartner M., Erdel M., Lee J.N., Paik Y.-K.,
 RA Glossmann H., Utermann G., Moebius F.F.;
 RT "Mutations in the delta7-sterol reductase gene in patients with the
 RT Smith-Lemli-Opitz syndrome.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:8181-8186(1998).
 RN [7]
 RP VARIANT SLOS IIE-289.
 RX PubMed=10995508;
 RA Krakowiak P.A., Nokoro N.A., Wassif C.A., Battaille K.P.,
 RA Nowaczyk M.J.M., Connor W.E., Maslen C., Steiner R.D., Porter F.D.;
 RT "Mutation analysis and description of sixteen RSH/Smith-Lemli-Opitz
 RT syndrome patients: polymerase chain reaction-based assays to simplify
 RT genotyping.";
 RL Am. J. Med. Genet. 94:214-227(2000).
 RN [8]
 RP VARIANTS SLOS MET-93; LEU-326; TRP-352 AND CYS-404.
 RX PubMed=11175299;
 RA Witsch-Baumgartner M., Clara E., Löffler J., Menzel H.J., Seedorf U.,
 RA Burn J., Gilleßen-Kaesbach G., Hoffmann G.F., Fitzky B.U., Mundy H.,
 RA Clayton P., Kelley R.I., Krajewska-Watasek M., Utermann G.,
 RT "Frequency gradients of DHCR7 mutations in patients with Smith-Lemli-
 RT Opitz syndrome in Europe: evidence for different origins of common
 RT mutations.";
 RL Eur. J. Hum. Genet. 9:45-50(2001).
 CC -1- FUNCTION: Production of cholesterol by reduction of C7-C8 double
 CC bond of 7-dehydrocholesterol (7-DHC).
 CC -1- CATALYTIC ACTIVITY: Cholesterol + NADPH(+) -> cholesta-5,7-dien-3-
 CC beta-ol + NADPH.
 CC -1- PATHWAY: Cholesterol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
 CC reticulum.
 CC -1- TISSUE SPECIFICITY: Most abundant in adrenal gland, liver, testis,
 CC and brain.
 CC -1- DISEASE: Defects in DHCR7 are the cause of Smith-Lemli-Opitz
 CC syndrome (SLOS or SLO). SLOS is a frequent inborn disorder of
 CC sterol metabolism with characteristic congenital malformations and
 CC dysmorphias. All patients suffer from mental retardation. Children
 CC with SLOS have elevated serum 7-dehydrocholesterol (7-DHC) levels
 CC and low serum cholesterol levels. SLOS occurs in relatively high
 CC frequency: approximately 1 in 20,000 to 30,000 births in
 CC populations of northern and central European background.
 CC -1- SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: AF096305; AAD09766.1; -
 CC EMBL: AF034544; AAC05086.1; -
 CC EMBL: AF110060; AAD24762.1; -
 CC EMBL: AF067127; AAD02816.1; -
 CC EMBL: BC000054; AAH00054.1; -
 CC EMBL: AF062481; AAC18345.1; -
 CC Genew: HGNC:2860; DHCR7.
 CC MIM: 602858; -
 CC MIM: 270400; -

```

DR MIM: 268670: -.
DR InterPro: IPR001171: ERG4_ERG24.
DR Pfam: PF01222: ERG4_ERG24_1.
DR PROSITE: PS01017; STEROL_REDUCT_1; 1.
DR PROSITE: PS01018; STEROL_REDUCT_2; 1.
DR Sterol biosynthesis; Cholesterol biosynthesis; Oxidoreductase; NADP;
KM Transmembrane; Endoplasmic reticulum; Disease mutation.
FT TRANSMEM 37 59
FT TRANSMEM 151 173
FT TRANSMEM 178 200
FT TRANSMEM 264 286
FT TRANSMEM 332 354
FT TRANSMEM 414 436
FT TRANSMEM 51 51
FT VARIANT 93 93
FT VARIANT 99 99
FT VARIANT 119 119
FT VARIANT 157 157
FT VARIANT 244 244
FT VARIANT 247 247
FT VARIANT 248 248
FT VARIANT 289 289
FT VARIANT 326 326
FT VARIANT 352 352
FT VARIANT 380 380
FT VARIANT 404 404
FT VARIANT 410 410
FT CONFLICT 5 5
FT CONFLICT 14 14
SQ SEQUENCE 475 AA; 54489 MW; 7D726443834C4EEB CRC64;

Query Match 26.9%; Score 698.5; DB 1; Length 475;
Best Local Similarity 35.3%; Pred. No. 2.6e-44;
Matches 154; Conservative 87; Mismatches 152; Indels 43; Gaps 10;

QY 13 ASMLSLAFPPVILMTYMHQD-----GSVTQTEGFEMWNGVGLINIM--PPR 62
DB 41 ASVIFLLFAP---FIVYYFIMACDDQYSCALIGPVVDI-----TGHARLSDIMAKTPPI 92
QY 63 TLIAWKIIFCYGAPALILLP-----GKRVGSPISPAENRPPYKANGLAAYF 111
DB 93 TRRAAOLYTLTMFGVLYLSLDPFCHKLPYGVGSIQEAIVPAQVYKNGYQINGQAWL 152
QY 112 VTLATYTLGLWLF-----GIENPAIVYDH-TGEIFSLIFGSIIFCVLLYTIKGHPASS 164
DB 153 LT-----HLIMFANAHLLSWFSPTIIFDWTPLIKCANILG-YAVSTFPAWKKYFFPTTA 206
QY 165 -DSGSGCNLLIIEYWNELYPRIKSGSDIKVFTNCRGMSMAVLAIVTYCIKOYEINGKV 223
DB 207 RCKKFGNFYNYMGMIEFNPRIGKMFDFKLFNGRGPVIAWTLINLSFAAKQRELHSHV 266
QY 224 SDSMLVNTIIMLVYVTKFEMWEGYNTNDIAHRCGFYICWGLVWVSVTSPMTYV 283
DB 267 TNAVIVANNVLAIVYIDFENWETWYLTIDICDHGFWIGWDCVWLPYLYTLQGLYLV 326
QY 284 NHVEVETGLAIVYIIVAGLICITINDDCROREPRRTNGKCLYMGARAPSKIYASTTTS 343
DB 327 YHPOVSTPHAVGVLLGLGVYIIFRYVANHOKDLFRRTDGRCLIMGKRKPKVLECYTSAD 386

```

```

QY 344 GETKTSLLITSGMWGLARHHEHYVPEILSAFEWYVPALEFNFVLYFYIFLLIFDAKR 403
DB 387 GORHRSKILVSGFVGVARHHEHYVVDIMGSLAYCAGGGHLLPRFYIYMAILLTTECLR 446
QY 404 DDDRCKSYKGYKWLK 419
DB 447 DEHRCASRYGDSWERY 462

RESULT 4
LBR_CHICK STANDARD; PRT; 637 AA.
AC P23913;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lamin B receptor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91009487; PubMed=2170422;
RA "Worman H.J., Evans C.D., Blobel G.;
RT "The lamin B receptor of the nuclear envelope inner membrane: a
RT polytopic protein with eight potential transmembrane domains.";
RL J. Cell Biol. 111:1535-1542(1990).
CC -1 FUNCTION: ANCHORS THE LAMINA AND THE HETEROCHROMATIN TO THE INNER
CC NUCLEAR MEMBRANE. CAN INTERACT WITH CHROMODOMAIN PROTEINS.
CC -1 SUBCELLULAR LOCATION: NUCLEAR ENVELOPE INNER MEMBRANE.
CC -1 PM: UNDERGOES PHOSPHORYLATION BY CDC2 PROTEIN KINASE IN MITOSIS
CC WHEN THE INNER NUCLEAR MEMBRANE BREAKS DOWN INTO VESICLES THAT
CC DISSOCIATE FROM THE LAMINA AND THE CHROMATIN. IT IS PHOSPHORYLATED
CC BY DIFFERENT PROTEIN KINASES IN INTERPHASE WHEN THE MEMBRANE IS
CC ASSOCIATED WITH THESE STRUCTURES. PHOSPHORYLATION OF LBR AND HPL
CC PROTEINS MAY BE RESPONSIBLE FOR SOME OF THE ALTERATIONS IN
CC CHROMATIN ORGANIZATION AND NUCLEAR STRUCTURE WHICH OCCUR AT
CC VARIOUS TIMES DURING THE CELL CYCLE.
CC -1 SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y00822; CA68758.1; -.
DR PIR: A36427; A36427.
DR InterPro: IPR001171; ERG4_ERG24.
DR InterPro: IPR002999; Tudor.
DR Pfam: PF01222; ERG4_ERG24_1.
DR SMART: SM00333; TUDOR; 1.
DR PROSITE: PS01017; STEROL_REDUCT_1; 1.
DR PROSITE: PS01018; STEROL_REDUCT_2; 1.
KM Receptor; Transmembrane; Phosphorylation; Nuclear protein;
KM DNA-binding.
FT DOMAIN 1 205
FT TRANSMEM 206 226
FT TRANSMEM 250 270
FT TRANSMEM 288 309
FT TRANSMEM 317 338
FT TRANSMEM 378 399
FT TRANSMEM 403 425
FT TRANSMEM 466 486
FT TRANSMEM 554 574
FT MOD_RES 95 95
FT MOD_RES 96 96
SQ SEQUENCE 637 AA; 73497 MW; 69D299002DBBCE0 CRC64;

```

```

RESULT 5
LBR_HUMAN
ID      LBR_HUMAN          STANDARD;          PRT;          615 AA.
AC      Q14739; Q14740;
DT      01-NOV-1997 (Rel. 35, Created)
DT      15-JUN-2002 (Rel. 41, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Lamin B receptor (Integral nuclear envelope inner membrane protein)
DE      (LMNB2).
GN      LBR.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94209307; PubMed=8157662;
RA      Ye Q., Worman H.J.;
RT      "Primary structure analysis and lamin B and DNA binding of human LBR."
RL      J. Biol. Chem. 269:11306-11311(1994).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94209308; PubMed=8157663;
RA      Schuler E., Lin F., Worman H.J.;
RT      "Characterization of the human gene encoding LBR, an integral protein
RT      of the nuclear envelope inner membrane."
RL      J. Biol. Chem. 269:11312-11317(1994).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      TISSUE=Skin;
RA      Strausberg R.;
RN      submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC      -1- FUNCTION: ANCHORS THE LAMINA AND THE HETEROCHROMATIN TO THE INNER
CC      NUCLEAR MEMBRANE. CAN INTERACT WITH CHROMOANIN PROTEINS.
CC      -1- SUBCELLULAR LOCATION: NUCLEAR ENVELOPE INNER MEMBRANE.
CC      -1- PFM: UNDERGOES PHOSPHORYLATION BY CDC2 PROTEIN KINASE IN MITOSIS
CC      WHEN THE INNER NUCLEAR MEMBRANE BREAKS DOWN INTO VESICLES THAT
CC      DISSOCIATE FROM THE LAMINA AND THE CHROMATIN. IT IS PHOSPHORYLATED

```

CC	DR	EMBL	L25931	AAA59495.1	1	-
DR <td>EMBL <th>L25941</th> <td>AAA59495.1 <td>1 <td>-</td> <td></td> </td></td></td>	EMBL <th>L25941</th> <td>AAA59495.1 <td>1 <td>-</td> <td></td> </td></td>	L25941	AAA59495.1 <td>1 <td>-</td> <td></td> </td>	1 <td>-</td> <td></td>	-	
DR <td>EMBL <th>L25932</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td></td>	EMBL <th>L25932</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td>	L25932	AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td>	1 <td>JOINED</td> <td></td>	JOINED	
DR <td>EMBL <th>L25933</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td></td>	EMBL <th>L25933</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td>	L25933	AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td>	1 <td>JOINED</td> <td></td>	JOINED	
DR <td>EMBL <th>L25934</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td></td>	EMBL <th>L25934</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td>	L25934	AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td>	1 <td>JOINED</td> <td></td>	JOINED	
DR <td>EMBL <th>L25935</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td></td>	EMBL <th>L25935</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td>	L25935	AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td>	1 <td>JOINED</td> <td></td>	JOINED	
DR <td>EMBL <th>L25936</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td></td>	EMBL <th>L25936</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td>	L25936	AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td>	1 <td>JOINED</td> <td></td>	JOINED	
DR <td>EMBL <th>L25937</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td></td>	EMBL <th>L25937</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td>	L25937	AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td>	1 <td>JOINED</td> <td></td>	JOINED	
DR <td>EMBL <th>L25938</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td></td>	EMBL <th>L25938</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td>	L25938	AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td>	1 <td>JOINED</td> <td></td>	JOINED	
DR <td>EMBL <th>L25939</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td></td>	EMBL <th>L25939</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td>	L25939	AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td>	1 <td>JOINED</td> <td></td>	JOINED	
DR <td>EMBL <th>L25940</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td></td>	EMBL <th>L25940</th> <td>AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td></td>	L25940	AAA59495.1 <td>1 <td>JOINED</td> <td></td> </td>	1 <td>JOINED</td> <td></td>	JOINED	
DR <td>EMBL <th>BC020079</th> <td>AAH20079.1 <td>1 <td>-</td> <td></td> </td></td></td>	EMBL <th>BC020079</th> <td>AAH20079.1 <td>1 <td>-</td> <td></td> </td></td>	BC020079	AAH20079.1 <td>1 <td>-</td> <td></td> </td>	1 <td>-</td> <td></td>	-	
DR <td>Genew <th>HGNC</th> <td>6518 <td>1 <td>LBH</td> <td></td> </td></td></td>	Genew <th>HGNC</th> <td>6518 <td>1 <td>LBH</td> <td></td> </td></td>	HGNC	6518 <td>1 <td>LBH</td> <td></td> </td>	1 <td>LBH</td> <td></td>	LBH	
DR <td>MIM <th>600024</th> <td>-</td> <td>-</td> <td></td> <td></td> </td>	MIM <th>600024</th> <td>-</td> <td>-</td> <td></td> <td></td>	600024	-	-		
DR <td>InterPro <th>IPR001171</th> <td>ERGA_ERG24</td> <td></td> <td></td> <td></td> </td>	InterPro <th>IPR001171</th> <td>ERGA_ERG24</td> <td></td> <td></td> <td></td>	IPR001171	ERGA_ERG24			
DR <td>InterPro <th>IPR002899</th> <td>Tudor</td> <td></td> <td></td> <td></td> </td>	InterPro <th>IPR002899</th> <td>Tudor</td> <td></td> <td></td> <td></td>	IPR002899	Tudor			
DR <td>Pfam <th>PF01222</th> <td>ERGA_ERG24</td> <td>1</td> <td></td> <td></td> </td>	Pfam <th>PF01222</th> <td>ERGA_ERG24</td> <td>1</td> <td></td> <td></td>	PF01222	ERGA_ERG24	1		
DR <td>SMART <th>SM00333</th> <td>TUDOR</td> <td>1</td> <td></td> <td></td> </td>	SMART <th>SM00333</th> <td>TUDOR</td> <td>1</td> <td></td> <td></td>	SM00333	TUDOR	1		
DR <td>PROSITE <th>PS01017</th> <td>STEROL_REDUCT_1 <td>1</td> <td></td> <td></td> </td></td>	PROSITE <th>PS01017</th> <td>STEROL_REDUCT_1 <td>1</td> <td></td> <td></td> </td>	PS01017	STEROL_REDUCT_1 <td>1</td> <td></td> <td></td>	1		
DR <td>PROSITE <th>PS01018</th> <td>STEROL_REDUCT_2 <td>1</td> <td></td> <td></td> </td></td>	PROSITE <th>PS01018</th> <td>STEROL_REDUCT_2 <td>1</td> <td></td> <td></td> </td>	PS01018	STEROL_REDUCT_2 <td>1</td> <td></td> <td></td>	1		
KW <td data-cs="6" data-kind="parent">Receptor; Transmembrane; Phosphorylation; Nuclear protein; DNA-binding.</td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td> <td data-kind="ghost"></td>	Receptor; Transmembrane; Phosphorylation; Nuclear protein; DNA-binding.					
KW <td>DOMAIN <td>1</td> <td>208</td> <td></td> <td></td> <td></td> </td>	DOMAIN <td>1</td> <td>208</td> <td></td> <td></td> <td></td>	1	208			
FT <td>TRANSMEM <td>212</td> <td>232</td> <td></td> <td></td> <td>NUCLEOPLASMIC (POTENTIAL).</td> </td>	TRANSMEM <td>212</td> <td>232</td> <td></td> <td></td> <td>NUCLEOPLASMIC (POTENTIAL).</td>	212	232			NUCLEOPLASMIC (POTENTIAL).
FT <td>TRANSMEM <td>258</td> <td>278</td> <td></td> <td></td> <td>POTENTIAL.</td> </td>	TRANSMEM <td>258</td> <td>278</td> <td></td> <td></td> <td>POTENTIAL.</td>	258	278			POTENTIAL.
FT <td>TRANSMEM <td>299</td> <td>319</td> <td></td> <td></td> <td>POTENTIAL.</td> </td>	TRANSMEM <td>299</td> <td>319</td> <td></td> <td></td> <td>POTENTIAL.</td>	299	319			POTENTIAL.
FT <td>TRANSMEM <td>326</td> <td>346</td> <td></td> <td></td> <td>POTENTIAL.</td> </td>	TRANSMEM <td>326</td> <td>346</td> <td></td> <td></td> <td>POTENTIAL.</td>	326	346			POTENTIAL.
FT <td>TRANSMEM <td>386</td> <td>406</td> <td></td> <td></td> <td>POTENTIAL.</td> </td>	TRANSMEM <td>386</td> <td>406</td> <td></td> <td></td> <td>POTENTIAL.</td>	386	406			POTENTIAL.
FT <td>TRANSMEM <td>447</td> <td>467</td> <td></td> <td></td> <td>POTENTIAL.</td> </td>	TRANSMEM <td>447</td> <td>467</td> <td></td> <td></td> <td>POTENTIAL.</td>	447	467			POTENTIAL.
FT <td>TRANSMEM <td>481</td> <td>501</td> <td></td> <td></td> <td>POTENTIAL.</td> </td>	TRANSMEM <td>481</td> <td>501</td> <td></td> <td></td> <td>POTENTIAL.</td>	481	501			POTENTIAL.
FT <td>TRANSMEM <td>561</td> <td>581</td> <td></td> <td></td> <td>POTENTIAL.</td> </td>	TRANSMEM <td>561</td> <td>581</td> <td></td> <td></td> <td>POTENTIAL.</td>	561	581			POTENTIAL.
FT <td>CONFLICT <td>154</td> <td>154</td> <td></td> <td></td> <td>S -> N (IN REF. 2).</td> </td>	CONFLICT <td>154</td> <td>154</td> <td></td> <td></td> <td>S -> N (IN REF. 2).</td>	154	154			S -> N (IN REF. 2).
FT <td>CONFLICT <td>301</td> <td>301</td> <td></td> <td></td> <td>A -> P (IN REF. 1).</td> </td>	CONFLICT <td>301</td> <td>301</td> <td></td> <td></td> <td>A -> P (IN REF. 1).</td>	301	301			A -> P (IN REF. 1).
FT <td>CONFLICT <td>530</td> <td>530</td> <td></td> <td></td> <td>T -> S (IN REF. 1).</td> </td>	CONFLICT <td>530</td> <td>530</td> <td></td> <td></td> <td>T -> S (IN REF. 1).</td>	530	530			T -> S (IN REF. 1).
QO <td>SEQUENCE <td>615 AA</td> <td>70703 MW</td> <td>5A7388776F43C66D</td> <td>CRC64</td> <td></td> </td>	SEQUENCE <td>615 AA</td> <td>70703 MW</td> <td>5A7388776F43C66D</td> <td>CRC64</td> <td></td>	615 AA	70703 MW	5A7388776F43C66D	CRC64	

Query Match 1: 21.0%; Score 544; DB 1; Length 615;
Best Local Similarity 32.9%; Pred. No. 7.4e-33;
Matches 143; Conservative 65; Mismatches 156; Indels 70; Gaps 16;

Db 424 NS-FOLLVVDALMNEALLTMDIHDGFMALGDLVWPFIYSFOAFYLVSHPNV 482
 QY 290 GFOALAIYIIVAGILCIYIN-DCDROROEFRRTNGKCLVWGRAPSKIVASYTTTSETKT 348
 Db 483 SPMASALITIVLK-LCGVIFRGANSOKNAFRN-----PSPKLAHLKTHITSTG 531
 QY 349 SLLTSGMGLARHFFHYVPEILSAFFWTPALPDNFLAYFYVILTLFLFRARDDDR 408
 Db 532 KNLVSGMGVGFHNPVLDGLIMALAMSLPCGFNHLPEFYIILFVMLVHREARDEYHC 591
 QY 409 RSKYGYKWKIYCEK 422
 Db 592 KKKYGVAMEKYYCOR 605

RESULT 6

ER24_ASCIM STANDARD: PRT: 430 AA.
 ID ER24_ASCIM STANDARD: PRT: 430 AA.
 AC P78575;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Delta(14)-sterol reductase (EC 1.3.1.70) (C-14 sterol reductase)
 GN ERG3.
 OS Ascombolus immerus.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Pezizomycetes;
 OC Pezizales; Ascombolaceae; Ascombolus.
 OX NCBI_TaxID=5191;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RN42;
 RA Kasbekar D.P.;

RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REDUCES THE C14-C15 DOUBLE BOND OF 4,4-DIMETHYL-
 CC CHOLESTA-8,14,24-TRIENOL TO PRODUCE 4,4-DIMETHYL-
 CC CHOLESTA-8,14,24-TRIENOL (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 4,4-dimethyl-5-alpha-cholesta-8,24-dien-3-
 CC beta-ol + NADP(+) = 4,4-dimethyl-5-alpha-cholesta-8,14,24-trien-3-
 CC beta-ol + NADPH.
 CC -1- PATHWAY: Ergosterol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

DR EMBL: Y10624; CAA17650.1;
 DR InterPro: IPR001171; ERG4_ERG24.
 DR Pfam: PF01222; ERG4_ERG24; 1.
 DR PROSITE: PS01017; STEROL_REDUCT_1; 1.
 DR PROSITE: PS01018; STEROL_REDUCT_2; 1.
 KW Sterol biosynthesis; Oxidoreductase; NADP; Transmembrane.
 FT TRANSMEM 12 32 POTENTIAL.
 FT TRANSMEM 67 89 POTENTIAL.
 FT TRANSMEM 109 129 POTENTIAL.
 FT TRANSMEM 230 250 POTENTIAL.
 FT TRANSMEM 267 287 POTENTIAL.
 FT TRANSMEM 290 310 POTENTIAL.
 FT TRANSMEM 349 369 POTENTIAL.
 FT TRANSMEM 376 396 POTENTIAL.
 SQ SEQUENCE 430 AA; 49126 MW; 7A61F01A4701B59D CRC64;

Query Match 19.4%; Score 503.5; DB 1; Length 430;
 Best Local Similarity 31.7%; Pred. No. 5e-30;
 Matches 140; Conservative 67; Mismatches 175; Indels 59; Gaps 15;

QY 8 PIVTASMLSLAFPCPFVILLWYTWVHODGSGVTFGFGEWGVGLINIMPR-----PT 63

Db 11 PIGTGVLMIL-----PISHYHLFLITPRGAPPE-----FWSAPLETLKSVTPFSSLS 62
 QY 64 LIAMKIIFCYGAFAEIIIDLLPGKRVEGPISPAGNRPYKANGLAAYVATATLG----- 119
 Db 63 LHATLAVAAYVLLVVALMYLPAELAEGLVLDKDSRLKRYCNAFTT-FLVEFFLGMTV 121
 QY 120 ----LWFGINFPNAVYHLEIFSFALIFGIFCVLLYIGH-VAPSS-----SDSGSC 169
 Db 122 LEGPTWFMWST-----LTDNFAQLOSASIVFYSALSWIYISYPMKRGKVILSPGFK 177
 QY 170 GNLLIDFYGMELYPRIKSFIDIVFTNCRFGMSMAVLAATYCIKOYEINGKYSMDLV 229
 Db 178 GNHIDFWMGRELNPRIEIMDIKOLHELRLGLMGWILFNLAMTVKQYNTGHFVSDSIVL 237
 QY 230 NTIIMLVYVTFEFWWEAGVWNTDIADHGRGFYICMGCLVWVSPVYISPGKYLNVHVEL 289
 Db 238 VNLFEPTWVVDALMNEKSVLTMDITDGLVMLFLGNAVWPPVYCLQARYLASFPVHL 297
 QY 290 GFOALAIYIIVAGILCI-YIN-----DCDROROEFRRTNGKCLVWGRAPSKIVASY-----T 340
 Db 298 G-----LGIAGVLAVOPTGTAIRGANNOXNAF-RTN-----PADPAVSHLFTMT 342
 QY 341 TTSGFTKSTLLTSGMGLARHFFHYVPEILSAFFWTPALPDNFLAYFYVILTLFLDR 400
 Db 343 TKSG-----SKLLISGMGVARHNVFGDMIMWAGYCLTGTGNTFLTYFYVYFGILLHR 398
 QY 401 AKRDDDRCKRYGKWKIYCE 421
 Db 399 DRDRKACRERYGKMDWDRYCK 419

RESULT 7

ER24_SCHPO STANDARD: PRT: 424 AA.
 ID ER24_SCHPO STANDARD: PRT: 424 AA.
 AC Q09195;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Delta(14)-sterol reductase (EC 1.3.1.70) (C-14 sterol reductase)
 GN ERG24 OR SPBC1655.18.
 GN Schizosaccharomyces pombe (Fission yeast).
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomyces.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=SP66;
 RA MEDLINE=95212923; PubMed=7698661;
 RA Smith S.;
 RT "Cloning and sequence analysis of an ERG24 homolog from
 RT Schizosaccharomyces pombe.";
 RL Gene 155:139-140(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltham J., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris J., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagsels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeli C.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds R.G., Squares R., Stevens S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voiclaert G., Aert R., Roben J., Gymnopoulos B.,
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Filtz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Wambut R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaur V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rocher M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McComb W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.,
RT "The genome sequence of *Schizosaccharomyces pombe*.";
RL Mature 415:871-880(2002).
CC -1- FUNCTION: REDUCES THE C14=C15 DOUBLE BOND OF 4,4-DIMETHYL-
CC CHOLESTA-8,14,24-TRIENOL TO PRODUCE 4,4-DIMETHYL-CHOLESTA-8,24-
CC DIENOL (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 4,4-dimethyl-5-alpha-cholesta-8,24-dien-3-
CC beta-ol + NADPH(+) = 4,4-dimethyl-5-alpha-cholesta-8,14,24-trien-3-
CC beta-ol + NADPH.
CC -1- PATHWAY: Ergosterol biosynthesis.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L36039; AAA121.1; -;
DR EMBL; AL023554; CAA19037.1; -;
DR InterPro: IPR001171; ERG4.ERG24.
DR Pfam: PF01222; ERG4.ERG24: 1
DR PROSITE: PS01017; STEROL_REDUCT_1; 1.
DR PROSITE: PS01018; STEROL_REDUCT_2; 1.
DR Sterol biosynthesis; Oxidoreductase; NADP: Transmembrane.
FT TRANSMEM 19 39 POTENTIAL.
FT TRANSMEM 112 132 POTENTIAL.
FT TRANSMEM 370 390 POTENTIAL.
SQ SEQUENCE 424 AA; 48560 MW; 1D3CE704F1B5E5B CRC64;

Query Match 19.3%; Score 500.5; DB 1; Length 424;
Best Local Similarity 32.5%; Pred. No. 8,1e-30;
Matches 128; Conservative 67; Mismatches 172; Indels 27; Gaps 11;

QY 43 TFGFENGNGV-----LIMIPRPTLAMKIFCYGAFALDQL--LGKRYEG- 91
DB 32 SFGSFTICNEGGCPAKFSKISHFKTPFLDQKSLILYLLMFSTLLMLKCTGNKAKGT 91
QY 92 PISPAGNRPPYKANGL-AAFVTLATYGLMWFGEFNPDAIVDH-LGEISALIFGSEIF 149
DB 92 PIDDKGRRLLYKINGFSACLLIGVCTSYLLGASCMERINDFQLMAAIVF-SVYL 150
QY 150 CVLLYIKGHVAPSS-SDSGSCGNIIDFYWGMELYPRIKGFIDIKYFNCRGMSMAVL 208
DB 151 CTFCYVQSFQKQDLAKGTSGNILFDWFGISLNPRIQ-NPDIKFCFCELRPGILIMVVE 209
QY 209 AVTYICIKQY-EINGKVSMSLVNTIIMLVYTFEWMGAGYNTMDIAHGRGFYICWG 267
DB 210 DIFACQIYLVGRITDSWLVLIIFHTWVDSLINESAVLTMTDTTGDGFYISFG 269
QY 268 LWWVPVSYTPGMYLVNHPPELTQALAYLLVAGIICITVINYDCDRORFRRNGKCIY 327
DB 270 LWWVPVSYTPGMYLVNHPPELTQALAYLLVAGIICITVINYDCDRORFRRNGKCIY 326
QY 328 WGRAPSVIYATYTTSETTSLLISGWMGLARHHPYELLISAFWVYALPDLNFAY 387
DB 327 ----PKLKHUKFIOTKRGTR--LLISGWMGMARHINVEGDMIMAWMCLPAGFGSIPY 379
QY 388 FYVYFLFLFLFDRAKRDGCRSKYGYKMYLCE 421
DB 380 FYVYFVGLLVHNRARDHRCRKVYGEDMEKYCK 413

RESULT 8
ID ER24_HUMAN STANDARD; PRT; 418 AA.
AC 076062; 095982; 096621; 09664;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Delta(14)-sterol reductase (EC 1.3.1.70) (C-14 sterol reductase)
DE (Sterol C14-reductase) (Delta14-SR) (Transmembrane 7 superfamily
DE member 2) (Another new gene 1) (Putative sterol reductase SR-1).
GN TW7SF2 OR ANGI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP MEDLINE=98277456; PubMed=9615229;
RX Lemmens I.H., Kas K., Merregaert J., Van De Ven W.J.M.;
RA "Identification and molecular characterization of TW7SF2 in the FAUNA
RT gene cluster on human chromosome 11q13.";
RL Genomics 49:437-442(1998).
RN [2]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=9907347; PubMed=9878250;
RA Holmer L., Pezhman A., Worman H.J.;
RT "The human lamin B receptor/sterol reductase multigene family.";
RL Genomics 54:469-476(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Brain, and Eye;
RA Strausberg R.;
RT Submitted (May-2001) to the EMBL/Genbank/DBJ databases.
RN [4]
RP PubMed=11784322;
RX Roberti R., Benatti A.M., Galli G., Caruso D., Maras B., Aisa C.,
RA Beccari T., Della Fazio M.A., Servillo G.;
RT "Cloning and expression of sterol Delta14-reductase from bovine
RT liver.";
RL Eur. J. Biochem. 269:283-290(2002).
CC -1- FUNCTION: Involved in the conversion of lanosterol to cholesterol.
CC -1- CATALYTIC ACTIVITY: 4,4-dimethyl-5-alpha-cholesta-8,24-dien-3-
CC beta-ol + NADPH(+) = 4,4-dimethyl-5-alpha-cholesta-8,14,24-trien-3-
CC beta-ol + NADPH.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Endoplasmic
CC reticulum.
CC -1- TISSUE SPECIFICITY: Expressed in adult heart, brain, pancreas,
CC lung, liver, skeletal muscle, kidney, ovary, prostate, and testis,
CC but not detected in placenta, spleen, thymus, small intestine,
CC colon (mucosal lining), or peripheral blood leukocytes.
CC -1- SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a number of
CC sequencing problems as reported in Ref.2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF048704; AAC21457.1; ALT_FRAME.
DR EMBL; AF023676; AAC21450.1; ALT_FRAME.
DR EMBL; AF096303; AAD09769.1; -;
DR EMBL; AF096304; AAD09765.1; -;
DR EMBL; BC009052; AAD09052.1; -;
DR EMBL; BC012857; AAH12857.1; -;
DR EMBL; BC012857; AAH12857.1; -;
DR Genew; HGNC:11863; TW7SF2.
DR MIM; 603414; -;

DR InterPro: IPR001171; ERG4_ERG24.
 DR Pfam: PF01222; ERG4_ERG24; 1.
 DR PROSITE: PS01017; STEROL_REDUCT_1; 1.
 DR PROSITE: PS01018; STEROL_REDUCT_2; 1.
 DR Sterol biosynthesis; Oxidoreductase; NADP; Transmembrane;
 KW Endoplasmic reticulum; Polymorphism.
 FT TRANSMEM 13 35 POTENTIAL.
 FT TRANSMEM 62 81 POTENTIAL.
 FT TRANSMEM 102 124 POTENTIAL.
 FT TRANSMEM 129 148 POTENTIAL.
 FT TRANSMEM 255 277 POTENTIAL.
 FT TRANSMEM 287 304 POTENTIAL.
 FT TRANSMEM 355 377 POTENTIAL.
 FT VARIANT 299 299 I->T.
 FT FTID=VAR.012716.
 FT CONFLICT 179 179 L->V (IN REF. 3: AAH12857).
 SO SEQUENCE 418 AA; 46417 MW; 357C8ABE2BED918 CRC64;
 Query Match 19.3%; Score 499.5; DB 1; Length 418;
 Best Local Similarity 32.8%; Pred. No. 9.5e-30;
 Matches 123; Conservative 60; Mismatches 167; Indels 25; Gaps 8;
 QY 54 GLINTPRPTLAWKIIFCYGAFAEALQLLLPGKRVGSPISAGNRPYKANGLAAYFT 113
 DB 53 GLEVLMSPRALLIM--LAMLGLQALYLLPARKVAEGGELKDKSRRLRPINGFQALVLT 109
 QY 114 LNTYGLMWFGLFNPAYVDHGEISALIF---GSFTCVLLTYIKGVAPSS--SDSG 167
 DB 110 -ALLVGLG---MSAGLPGLALPEMLPLAFYATLTAFTFSFLYMKAOVAVSALAERG 164
 QY 168 SCGNLIIDFYWGMELYPRIKGSFDIKVTNCRFGMMASNAVAVYCIKOYEINGKVSQM 227
 DB 165 NSGNPIYDFELGRELNPRI-CEPDFKFCFELRBLIGVLLINLALMKRAELRGSPSLAM 223
 QY 228 LVNTIIMLYVYKFFWMEGAYNTMDIAHNRGFTYICAGCLVWVSVTSPGMYLVNHPV 287
 DB 224 WLVNGFOLLVYGDALMHEBAVLTTMDITHDGCFMLAFGDMAWVPFTYSLQAQFLHHPQ 283
 QY 288 ELGTOLATYIVAGILCIYINVDCCROBOEPRFTNGCKLVWGRAPSKIVASTYTTSGETK 347
 DB 284 PLGLPMASTICLINAIGYITFGANSOKMTFRKNPSDPRVAG-----LETSTATGKR- 336
 QY 348 TSLILTSGMWGLARHFRHYVPELISAFETVPALFDNFLAYFYVIELTLLLPDRAKRDOR 407
 DB 337 ---LWVGMMGWVRHNPVYGLDLMALWAGSLPGVSHLPEYILVYLTALLVREARDEKQ 393
 QY 408 CRKTYGKMYKLYCEK 422
 DB 394 CLKYGLAMOEYCR 408
 RESULT 9
 ER24_NEUCR STANDARD: PRT: 490 AA.
 AC P38670;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Delta(14)-sterol reductase (EC 1.3.1.70) (C-14 sterol reductase)
 DE (Sterol C14-reductase).
 GN ERG-3 OR ERG3.
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=74-OR23-1A;
 RA Papavinasandaram K.G., Kasbekar D.P.;
 RT "The Neurospora crassa erg3 gene encodes a protein with sequence
 RT homology to both yeast sterol C-14 reductase and chicken lamin B
 RT receptor.";
 RL J. Genet. 73:33-41(1994).

CC -1- FUNCTION: REDUCES THE C14=C15 DOUBLE BOND OF 4,4-DIMETHYL-
 CC CHOLESTA-8,14,24-TRIENOL TO PRODUCE 4,4-DIMETHYL-CHOLESTA-8,24-
 CC DIENOL.
 CC -1- CATALYTIC ACTIVITY: 4,4-dimethyl-5-alpha-cholesta-8,24-dien-3-
 CC beta-ol + NADP(+) = 4,4-dimethyl-5-alpha-cholesta-8,14,24-trien-3-
 CC beta-ol + NADPH.
 CC -1- PATHWAY: Ergosterol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X7795; CA54919.1; -.
 DR PIR: S44170; S44170.
 DR InterPro: IPR001171; ERG4_ERG24.
 DR Pfam: PF01222; ERG4_ERG24; 1.
 DR PROSITE: PS01017; STEROL_REDUCT_1; 1.
 DR PROSITE: PS01018; STEROL_REDUCT_2; 1.
 KW Sterol biosynthesis; Oxidoreductase; NADP; Transmembrane.
 FT TRANSMEM 23 43 POTENTIAL.
 FT TRANSMEM 80 100 POTENTIAL.
 FT TRANSMEM 136 156 POTENTIAL.
 FT TRANSMEM 160 180 POTENTIAL.
 FT TRANSMEM 230 250 POTENTIAL.
 FT TRANSMEM 255 275 POTENTIAL.
 FT TRANSMEM 324 344 POTENTIAL.
 FT TRANSMEM 436 456 POTENTIAL.
 SO SEQUENCE 490 AA; 54722 MW; 9CF3104EF3014085 CRC64;
 Query Match 18.3%; Score 473.5; DB 1; Length 490;
 Best Local Similarity 30.9%; Pred. No. 9e-28;
 Matches 133; Conservative 61; Mismatches 151; Indels 85; Gaps 18;
 QY 48 W-ENGVOGLINMPRLTIAWKIIFCYGAFAEALQLLLPGKRVGSPISAGNRPYKANG 106
 DB 75 WPDNGVGLGS-W--SATLWTL--GYALSLVYRFLPGHHVGTSLSGGRKLYKLA 128
 QY 107 LAAYVTLATYLG-----LMWGFNPAYVDHGEISALIFGFTCVLLTYIK 156
 DB 129 FNSAMCTLAIIAAGTIAOGAEFPVWTF-----ISDNFAQIISANILFAFALAIIFYVR 181
 QY 157 G-HVAPSSD-----SGSCNLIIDFYWGMELYPRIG---KSFDIKVTNCRFGMSW 205
 DB 182 SFDPVPGNKDMRQLAAGGVGSLIYDFYIGRELNPRIITPLIGOVDTKEFMENPGILGW 241
 QY 206 AVLAVTYCIRKYEINGKVSQSMNVNTIIMLYVYKFFWMEGAYNTMDIAHNRGFTYICW 265
 DB 242 IILNCAFIARQYRLXGYVTDLSILFTIAQAFYVDGIYMEVAVLTTMDITTDGCFMLSF 301
 QY 266 GCLVWVPSVYSPGMYLVNHPVELGT--QLAI-YLVAGILCIYINVDCCROBOEPRFTN 322
 DB 302 GDVWVWPPMYSTQRYLSVHPQAGAFGLAVGAVLAAGSIIFRLS--NSQKNF--RTN 357
 QY 323 GKCLVWGAAPSKIVASTYTTSGETKT-SLLTSGMWGLARHFRHYVPELISAFETVP--- 378
 DB 358 PE-----DPSVKHLTYL---QRTKTSRLITSGMWGIARHINVLGWLDSWPSLPTGI 407
 QY 379 -----ALFDN-----FLAFYVIELTLLLPDRAKRDOR 409
 DB 408 AGYQLLSAGSNAPGALITMLDREVYVGARWGIVFTYITLYIAILLIHRDLKDDKCS 467
 QY 410 SKYGYKMYKLY 419
 DB 468 KKYGDWMEKY 477
 RESULT 10

CC	Eukaryota;Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
CC	Hypocreales; Nectriaceae; Nectria.
OX	NCBI_TaxID=70791;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-MP VI;
RA	Srinivas G., Vanetten H.D., Kasbekar D.P.;
RL	Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
CC	-1- FUNCTION: REDUCES THE C14-C15 DOUBLE BOND OF 4,4-DIMETHYL-CHOLESTA-8,14,24-TRIENOL TO PRODUCE 4,4-DIMETHYL-CHOLESTA-8,24-DIENOL.
CC	-1- CATALYTIC ACTIVITY: 4,4-dimethyl-5-alpha-cholesta-8,24-dien-3-beta-ol + NADPH (+) = 4,4-dimethyl-5-alpha-cholesta-8,14,24-trien-3-beta-ol + NADPH.
CC	-1- PATHWAY: Ergosterol biosynthesis.
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC	-1- SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib.ch).
CC	-----
DR	EMBL; X94315; CAA63976.1; -
DR	InterPro; IPR001171; ERG4_ERG24.
DR	Pfam; PF01222; ERG4_ERG24_1.
DR	PROSITE; PS01017; STEROL_REDUCT_1; 1.
DR	PROSITE; PS01018; STEROL_REDUCT_2; 1.
KW	Sterol biosynthesis; Oxidoreductase; NADP; Transmembrane.
FT	TRANSMEM 18 38 POTENTIAL.
FT	TRANSMEM 77 97 POTENTIAL.
FT	TRANSMEM 131 151 POTENTIAL.
FT	TRANSMEM 155 175 POTENTIAL.
FT	TRANSMEM 319 339 POTENTIAL.
FT	TRANSMEM 431 451 POTENTIAL.
SO	SEQUENCE 485 AA; 54597 MW; 45828AD154B02B6 CRC64;

```

Db      465  YGDEWEKY 472

||||:|:|
||||:|:|

RESULT 12
ER24_SEPLY STANDARD: PRT; 512 AA.
AC 013597;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Delta(14)-sterol reductase (EC 1.3.1.70) (C-14 sterol reductase)
DE (Sterol C14-reductase).
GN ERG3.
OS Septoria lycopersici (Tomato leaf spot fungus).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothyriomycetes incertae sedis;
OC Mycosphaerellaceae; mitosporie Mycosphaerellaceae; Septoria.
OX NCBI_TaxID=39703;
RN [1]
RP SEQUENCE FROM N.A.
RA Aparna K.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: REDUCES THE C14-C15 DOUBLE BOND OF 4,4-DIMETHYL-
CC CHOLESTA-8,14,24-TRIENOL TO PRODUCE 4,4-DIMETHYL-CHOLESTA-8,24-
CC DIENOL (BY SIMILARITY).
CC -! CATALYTIC ACTIVITY: 4,4-dimethyl-5-alpha-cholesta-8,24-dien-3-
CC beta-ol + NADPH(+) = 4,4-dimethyl-5-alpha-cholesta-8,14,24-trien-3-
CC beta-ol + NADPH.
CC -! PATHWAY: Ergosterol biosynthesis.
CC -! SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -! SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; Y14389; CAA74747.1; -;
DR InterPro; IPR001171; ERG4_ERG24.
DR Pfam; PF01222; ERG4_ERG24; 1.
DR PROSITE; PS01017; STEROL_REDUCT_1; 1.
DR PROSITE; PS01018; STEROL_REDUCT_2; 1.
FT Sterol biosynthesis; Oxidoreductase; NADP; Transmembrane.
FT TRANSMEM 27 47 POTENTIAL.
FT TRANSMEM 100 120 POTENTIAL.
FT TRANSMEM 140 160 POTENTIAL.
FT TRANSMEM 172 192 POTENTIAL.
FT TRANSMEM 242 262 POTENTIAL.
FT TRANSMEM 278 298 POTENTIAL.
FT TRANSMEM 324 344 POTENTIAL.
FT TRANSMEM 353 373 POTENTIAL.
FT TRANSMEM 418 438 POTENTIAL.
FT TRANSMEM 458 478 POTENTIAL.
SQ SEQUENCE 512 AA; 57578 MW; 169A756E60CDE963 CRC64;

```

[illegible]

```

Query Match 17.7%; Score 458.5; DB 1; Length 512;
Best Local Similarity 27.7%; Pred. No. 1.2e-26;
Matches 137; Conservative 70; Mismatches 178; Indels 109; Gaps 16;

QY      8 PIVTA--SMLSLAFCPEVILL-----WYRWVHODGSVTOFGFEWEN 50
      || || : : : || | | | | | | | | | | | | | | | |
Db      37 PIACYAFGFLCDVSGCPEPILLSBKLETPPLTSKNKPMQ--HALDTLAEVG--WP- 90

QY      51 GYQGLINIPRTLIAWKIIFCYGAFEATLJOLPGKRGCEIPSPAGNRPVYKANG--IAA 109
      || || || | | | | | | | | | | | | | | | | | | | |
Db      91 GMSGLIN-----TEAVLGFEFNYG--LSLLMWLLPRAHEVEGELRTGGGLKTRFNACLSA 144

QY     110 YEVTLATYLG-----LMMFGFIENPAIVVDHLGCEIRSAIFGSEFICVLLYIKGHVA 160
      : : | | | | | | | | | | | | | | | | | | | | |

```

Db 145 VTIFVACAGTIVRGPEQWTF-----INRNYQLTFLNIIIAVALIYYLKSEV 197
 QY 161 PSSSD-----SGSCGNLIIDFWGMEIYPRICKSF-----DIKVFNCNCFGMSNAVLA 209
 Db 198 KAGTEORELAAGSHSHILYDMWGRLENPRLTTPITPGEVDIKSFELPFGMGWLLD 257
 QY 210 VTTCIKOYEINGKVSMSLVNTIML-----VYVYKFFWMEAGYWNMT 252
 Db 258 LAFNAKQKSGYITDSMRKWTPLLLGIVLTHPVVITISQSYVVDALYMEPAITLM 317
 QY 253 DIANDRGFYICWGLVWDSVYTSPOGYLVNHPVELGTOIAIYILVAGILCIYINYCD 312
 Db 318 DLTTDGFEMISFEDLVWVPEIYSIQAKYLSVHPVALGPRYVALITLITQATGYXI----- 372
 QY 313 RQREFRRTNKKLVWGRARASKIAYSTTSGEKTSLLLTSGMWGLARHPHYPELISA 372
 Db 373 -----FRATNNDKNIFETPNPDKVAHLKXIETGTGRLLTGTGWGFAHINVLGWLMS 427
 QY 373 FFWTVPA-----LFDNF-----LAFYVYFLLTLLFDRAKRDDR 407
 Db 428 WSYCLPFLAAGYKLTPLTLPENSRVSTDMKAGIPITYFYMLYFAILLHRRDEAK 467
 QY 408 CRSYKGYKWLKYE 421
 Db 488 CRKRYGAHWEKCYQ 501

RESULT 13

ER24_YEAST STANDARD; PRT; 438 AA.

AC P32462;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Delta(14)-sterol reductase (EC 1.3.1.70) (C-14 sterol reductase)
 GN ERG24 OR YNL280C OR NO593.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93039674; PubMed=1418625;
 RA Lorenz R.T., Parks L.W.;
 RT "Cloning, sequencing, and disruption of the gene encoding sterol C-14 reductase in Saccharomyces cerevisiae.";
 RL DNA Cell Biol. 11:685-692(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94171077; PubMed=8125337;
 RA Lai M.H., Bird M., Pierson C.A., Alexander J.F., Goebel M.,
 RT Carter G.T., Kirsch D.R.;
 RT "The identification of a gene family in the Saccharomyces cerevisiae ergosterol biosynthesis pathway.";
 RL Gene 140:41-49(1994).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Messenguy F., Dubois E., Vierendeels F., Scherens B., Pierard A.,
 RL Glandsdorf N.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: REDUCES THE C14-C15 DOUBLE BOND OF 4,4-DIMETHYL-CHOLESTA-8,14,24-TRIENOL TO PRODUCE 4,4-DIMETHYL-CHOLESTA-8,24-DIENOL.
 CC -1- CATALYTIC ACTIVITY: 4,4-dimethyl-5-alpha-cholesta-8,24-dien-3-beta-ol + NADPH.
 CC -1- ENZYME REGULATION: INHIBITED BY THE MORPHOLINE ANTIFUNGAL DRUG FENPROPIMORPH.
 CC -1- PATHWAY: Ergosterol biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its contents is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-slb.ch/announce/or send an email to license@isb-slb.ch>).

CC EMBL: M99419; AAA18256.1;
 DR EMBL: S69420; AAB30203.1;
 DR EMBL: Z71556; CAA96192.1;
 DR PIR: S30769; S30769.
 DR SGD: S0005224; ERG24.
 DR InterPro: IPR001171; ERG4_ERG24.
 DR Pfam: PF01222; ERG4_ERG24.
 DR PROSITE: PS01017; STEROL_REDUCT_1;
 DR PROSITE: PS01018; STEROL_REDUCT_2;
 KW sterol biosynthesis; oxidoreductase; NADP; Transmembrane.
 FT TRANSMEM 14 30
 FT TRANSMEM 72 90 POTENTIAL.
 FT TRANSMEM 110 127 POTENTIAL.
 FT TRANSMEM 148 172 POTENTIAL.
 FT TRANSMEM 243 263 POTENTIAL.
 FT TRANSMEM 309 328 POTENTIAL.
 FT TRANSMEM 369 387 POTENTIAL.
 FT CONFLICT 253 253 L->S (IN REF. 2).
 SQ SEQUENCE 438 AA; 50615 MW; B936B04506C0F3 CRC64;

Query Match 16.4%; Score 424.5; DB 1; Length 438;
 Best Local Similarity 29.7%; Pred. No. 3.2e-24;

Matches 130; Conservative 74; Mismatches 160; Indels 73; Gaps 20;

QY 23 PPVITLMTWYHODGSYQTFGFF-----WENGVO-----GLINIMPRPTLIAMK 68
 Db 28 PVFTIIL-NQMIRPDIYFIK---GFQNPDIYELM-NGIKPLRYIGNRELW----- 73
 QY 69 IIFC-YGAFEAIIQLLLPGKRVGEIPAGNRPRPYKANGLAAYVTLATYIGLMWEGIF 126
 Db 74 TVYCLWYGLT-LVLDVITLGRVWKGVLQRLDGSRLSYKINGI-AMSTTVLVLAIRWKLTD 131
 QY 127 N-PAIVY---DHLGEIFSLIFGSEI-----FCVLLYIK---GNVAPSSSDSGSCG 170
 Db 132 GQLPELOIYENHVSICILISLFFLAYCYVASFIPILFEKKNNGKREKILAGMSG 191
 QY 171 NLIDFPGWMEIYPRICKSEIDIKVFTNCRFGMSNAVALYVICIKOYEINGKVSMSLVN 230
 Db 192 NIIVDFIETREINPRIG-PLDIKMFSELRPMMLLILNLSCHLHHYLTGKINDALVY 250
 QY 231 TILMLVYVTKFFWMEAGYWNMTDIANDRGSEFYICWGLVWDSVYTSPOGYLVNHPVELG 290
 Db 251 NPLQGYITFDGVNLNEGVLTMDDITTDGFGFMALAGDSLVPFTYSLQARYLSVPELG 310
 QY 291 TOLAIYILVAGILCI-----YINYDOROROFRRRTNCKLWGRAPSKIVASYTTTSGE 345
 Db 311 -----WVKVYGLAIIMFELGFHFHSANKKSEFRQ---GKL-----ENKLSIOTKRG- 354
 QY 346 TYSLLISGWMGLARHPHYPELISAFETWYPALEDFNLAYVITLTLFDRAKRDD 405
 Db 355 TK---LHCDGWWAKSOHINIFGDWILSLSWCATATFORPLTYVYSLYVATLLHROQDE 411
 QY 406 DRCRSYKGYKWLKYE 422
 Db 412 HKCRLYGEMWEEYERK 428

RESULT 14

ERG4_YEAST STANDARD; PRT; 473 AA.

AC P23340;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Delta(24(24(1)))sterol reductase (EC 1.3.1.71) (Sterol delta(24(28))-

QY	15	MLSLLACPPFVILLWTTWHODGSVT-QTFPFW-----ENGVOGLINMPR	61
Db	46	VIGMLIEPFLMYMWCAPFYHGKVALPRAGSMNHFKHLVOLYLENGI-----	96
QY	62	PFTLAKRIIFCYGAFEAIILOLLPGKRVEG-PISP-AGNRPVYKANGLAAYEVLATYLG	119
Db	97	PEKDWTNIFLFPMWFQITLYTYTLPGIWTKGQPLSHLKGNOLPYFCAMMTLVTTTIVLV	156
QY	120	LMEGFENPAIVYDHLGEIFSALIFCSFIIPCVLLYIKGHAVDSSSDSGCNLIIDFPWG	179
Db	157	LHFNLFRLVYIIDRFGRIMTCALISGFAPESILTY-WTLFISHDYHMRGTGNNHLDYDFMG	215
QY	180	MELYPRIGKSPFDIKVFINGCFGMASNAVALVYTICIOYELENKYSDSMLVNTIMLVYT	239
Db	216	APLPBRMG-IIDLKMFEEVRLPWFTLYFTYLGACLXOWEMETGYGVYPQLGVAALHMLTAN	274
QY	240	KFFWNEGVNYNMIDIAHDRCFYICWGLVWVSYSPTSGMLVNH-PVELG----TOLA	294
Db	275	ACANGELIYPTMDMAVERKGFMILMNLAGVYTYCHCTIYLYINDPEHYMSUTLYNS	334
QY	295	IYILVAGIICITYIND-CDROQEER-----RTNGKLWVGRAPSKIYAS--YTTTSCG	345
Db	335	LYVV---ILCAVYEFDTANAQNKAFAFRKMOSGDKTGRKTEPF--LPYQLINKPKYMWTS--	387
QY	346	TKTSLLTSGWMGLARFHVHPBELSAFFPTVVALNDNFATYVLEFLLLPDAKRD	405
Db	388	-NGSYLLIDEMYTLARKRIHYTAADWTOSLVWALSCEFNSEVPWFEPVFLVLIHRAFRQ	446
QY	406	DRCRSKYGKWKLXCE	421
Db	447	AKCKRKXGKDMDEYCK	462
 RESULT_15			
STSL_SCOP0			
ID	SFS1_SCOP0	STANDARD;	PRT; 453 AA.
AC	P36209; O13891;		
DT	01-JUN-1994 (Rel. 29, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Delta(24)(24(1))}-sterol reductase (EC 1.3.1.71) (Sterol delta(24)(28) reductase) (C-24(28) sterol reductase).		
CN	SFS1 OR SPAC2064.07C.		
OS	Schizosaccharomyces pombe (fission yeast).		
OC	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;		
OC	Schizosaccharomycetales; Schizosaccharomycetaceae;		
OX	NCHI_Taxid=4896;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=92329994; PubMed=1320960;		
RA	Shimanuki M., Goebel M., Yanagida M., Toda T.:		
RT	"Fission yeast stt1 gene encodes a protein similar to the chicken lamin B receptor and is implicated in pleiotropic drug-sensitivity,		
RT	divalent cation-sensitivity, and osmoregulation."		
RL	Mol. Biol. Cell 3:263-273(1992).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=972;		
RX	MEDLINE=21848401; PubMed=11659360;		
RA	Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,		
RA	Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,		
RA	Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,		
RA	Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,		
RA	Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,		
RA	Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels J.,		
RA	James K., Jones L., Jones M., Leather S., McDonald S., Mclean J.,		
RA	Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,		
RA	O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,		
RA	Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,		
RA	Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,		
RA	Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.		

RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Mueller-Auer S.,
 RA Borzys K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambolt R., Purnelle B.,
 RA Gofreau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galbert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Dargatzis A., Revuelta J.L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen J., Potashkin J.,
 RA Shpakovski G.V., Useery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of *Schizosaccharomyces pombe*.";
 RL Nature 415:871-880(2002).
 [3]
 RP CHARACTERIZATION, PubMed=8125337,
 RX MEDLINE=94171077; Alexander J.F., Goebel M.,
 RA Lai M.H., Bard M., Pierson C.A.,
 RA Carter G.T., Kirsch D.R.;
 RT "The identification of a gene family in the *Saccharomyces cerevisiae*
 RT ergosterol biosynthesis pathway.";
 RL Gene 140:41-49(1994).
 CC -1- CATALYTIC ACTIVITY: Ergosterol + NADP(+) = ergosta-
 CC 5,7,22,24(24(1))-tetraen-3-beta-ol + NADPH.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
 CC -1- SIMILARITY: BELONGS TO THE ERG4/ERG24 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X63549; CAA45113.1; -
 DR EMBL: Z98600; CAB1256.1; -
 DR PIR: A43765; A43765.
 DR InterPro: IPR001171; ERG4_ERG24.
 DR Pfam: PF01222; ERG4_ERG24; 1.
 DR PROSITE: PS01017; STEROL_REDUCT_1; 1.
 DR PROSITE: PS01018; STEROL_REDUCT_2; 1.
 KW Sterol biosynthesis; Oxidoreductase; NADP; Transmembrane.
 FT TRANSMEM 13 33 POTENTIAL.
 FT TRANSMEM 77 97 POTENTIAL.
 FT TRANSMEM 118 138 POTENTIAL.
 FT TRANSMEM 153 173 POTENTIAL.
 FT TRANSMEM 209 229 POTENTIAL.
 FT TRANSMEM 278 298 POTENTIAL.
 FT TRANSMEM 311 331 POTENTIAL.
 FT TRANSMEM 399 419 POTENTIAL.
 FT CONFLICT 412 412 C -> S (IN REF. 1).
 SQ SEQUENCE 453 AA; 52545 MW; 4740B6EB3BD27CF CRC64;

Query Match 14.1%; Score 364.5; DB 1; Length 453;
 Best Local Similarity 26.8%; Pred. No. 8.5e-20;
 Matches 117; Conservative 69; Mismatches 204; Indels 47; Gaps 12;

QY 16 LSLAFCEPPEVILMYTWVHODGSVTQTFGFF--WENGVOGL--INIMPRPTLIAMKI 69
 Db 20 LAIMTGFCPLMYLWACSKFNDSQIFKESFTIAGFQNFRLTGHYIYGAVPTRYAFIV 79
 QY 70 IFCGAFPAIIQLLPGRVKG-PISPAQN-RPYKANGLAAYVTLATYIGLWMEGIEFN 127
 Db 80 FWSFCIOAVMYLPLPGVITGLPKHNRNRLPYLCNAIWSFYTTIIVILAVLHVTFFP 139
 QY 128 PAIVYDHGEFFSALIFSGFFCYLYI-----KGHVAPSSSDSGCGNLIDF 176
 Db 140 ITTFIDMGFLMSVAIIITAFVCTFVLTGTLFDRLFDKPH-----RLSGNPITYDA 191
 QY 177 YWGMELEPRICKGSPDIKVFITNCRFGMMSAVL--AVYCIKQYEINGKVSDSMLVNTIL 233

Db 192 EMGACLNPRILGKLLDFKMFEEVR--IPWFLFFISGAARQETVGVSPQVLFVCIG 248
 QY 234 MLVYVTKFFWMEAGYNTNDIAHRCGFTYCWGLVWVSVYTSPPMYLVNHPVEL---G 290
 Db 249 HYLVAANCSKGEOLIVPTWMAVEKGFMLIFWNNAGVPTYSHCTLYLFSHDPVYNN 308
 QY 291 TOLATYLVAGIICITYNDCDRQROEPR---TNGKCLWGRAPSKIVAST---TTS 343
 Db 309 TQYTTGIYVLLCCYIIFDTGNGQKHFNRQIYGTVEVHRKTFPQLPMLIKNPTFCAN 368
 QY 344 GETKTSLLTSGWMLARHFRHVPETLSAFWVTPALFNFILAYFVIFLTLLFDRAR 403
 Db 369 GGT---LTSGWRYARKIHTYADEFGSLSMALITGFQSPLEFYFPCFFVVLVHRSR 424
 QY 404 DDDRCNSKTYKWKLYC 420
 Db 425 DIRCKAKYGADEDEYC 441

Search completed: January 14, 2003, 14:15:32
 Job time: 16 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 14:14:09 : Search time 22 Seconds
(without alignments)
2079.999 Million cell updates/sec

Title: US-09-817-774-31

Perfect score: 2594

Sequence: 1 MAETVHSPIVTYASMLSLA.....RAKTKMQDAIDILLICML 476

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR73:*
2: PIR1:*
3: PIR2:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2312	89.1	432	2	sterol delta7 redu
2	554	21.4	637	2	lamin B receptor -
3	553.5	21.3	620	2	lamin B receptor -
4	540	20.8	615	2	lamin B receptor -
5	500.5	19.3	424	2	probable C-14 ster
6	473.5	18.3	490	2	probable C-14 ster
7	424.5	16.4	438	2	probable C-14 ster
8	418.5	16.1	697	2	hypothetical prote
9	371.5	14.3	317	2	nuclear envelope m
10	371	14.3	473	2	probable transpor
11	363.5	14.0	453	2	stsl+ protein - fi
12	120.5	4.6	527	2	cytochrome-c oxida
13	118.5	4.6	527	2	cytochrome-c oxida
14	115.5	4.5	527	2	cytochrome-c oxida
15	108	4.2	599	2	probable membrane
16	106	4.1	354	2	opsin, rod ultra
17	106	4.1	474	2	NADH2 dehydrogenas
18	105	4.0	499	2	alignate o-acetyl
19	105	4.0	504	2	probable ammonium
20	104.5	4.0	1418	2	hypothetical prote
21	104	4.0	444	2	NADH2 dehydrogenas
22	104	4.0	550	1	afub protein homol
23	102	3.9	438	1	preprotein transpo
24	101.5	3.9	514	2	ammonium transpo
25	101.5	3.9	522	2	cytochrome-c oxida
26	101.5	3.9	598	2	probable membrane
27	101	3.9	494	2	probable permease
28	100.5	3.9	485	2	Not56-like protein
29	100	3.9	438	2	Not56-like protein

30	100	3.9	485	2	F86108	probable peptide t
31	100	3.9	485	2	H91267	probable peptide t
32	100	3.9	494	2	F86646	hypothetical prote
33	98.5	3.8	558	2	S08270	cytochrome-c oxida
34	98	3.8	393	2	C95343	hypothetical prote
35	98	3.8	479	2	F86285	F9L1.11 protein -
36	98	3.8	1515	2	T52081	MRP-like ABC trans
37	97.5	3.8	173	2	D58893	NADH2 dehydrogenas
38	97.5	3.8	416	2	G81258	serine transporter
39	97	3.7	443	2	F30010	NADH2 dehydrogenas
40	96.5	3.7	322	2	E84908	hypothetical prote
41	96.5	3.7	333	2	T02690	hypothetical prote
42	96.5	3.7	477	2	D82587	cationic amino aci
43	96.5	3.7	524	1	ODR21	cytochrome-c oxida
44	96.5	3.7	528	1	ODR21	cytochrome-c oxida
45	96.5	3.7	639	2	T30151	hypothetical prote

ALIGNMENTS

RESULT 1
F96540
sterol delta7 reductase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: F96540
R:Rheologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chlin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hultzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96540
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-432 <STO>
A:Cross-references: GB:AE005173; NID:g9454565; PID:AF87888.1; GSPDB:GN00141
A:Gene: F1P12.21
A:Map position: 1
C:Superfamily: yeast probable C-14 sterol reductase
Query Match 89.1%; Score 2312; DB 2; Length 432;
Best Local Similarity 99.8%; Pred. No. 4.2e-188;
Matches 421; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	MAETVHSPIVTYASMLSLAFCPPVILLIWMVHODSVTQTFGEFMEVNGVGLINTIP	60
DB	1	MAETVHSPIVTYASMLSLAFCPPVILLIWMVHODSVTQTFGEFMEVNGVGLINTIP	60
QY	61	RPTLLAMKTIIFCYGAFEAIIQLLPGRKREGPISNAGNRPYKANGLAAYTTATATYIGL	120
DB	61	RPTLLAMKTIIFCYGAFEAIIQLLPGRKREGPISNAGNRPYKANGLAAYTTATATYIGL	120
QY	121	WMFGFENPAIVDHDGEIFSAIFGFCVLLYKGVAPSSDGSAGNLIIDFYVGM	180
DB	121	WMFGFENPAIVDHDGEIFSAIFGFCVLLYKGVAPSSDGSAGNLIIDFYVGM	180
QY	181	EYPRIGKSFDFIKVTNCRFGMMSNAVLAVYCIKQYEINRGVSDSMVNTILMLVYTK	240
DB	181	EYPRIGKSFDFIKVTNCRFGMMSNAVLAVYCIKQYEINRGVSDSMVNTILMLVYTK	240
QY	241	FWMEGAVNTMDIADHGRGFTICGCLVWPVSYYTSGMILVNPBVLGQALATYIYA	300
DB	241	FWMEGAVNTMDIADHGRGFTICGCLVWPVSYYTSGMILVNPBVLGQALATYIYA	300
QY	301	GILCIYINDCDROROEFRRTNGKCLVWGRAFSKIIVASTYTTSGETKTSLLTSGMWGLA	360

```

Db 301 GLCIYINDDCRQROEFRTNKGKLVWRAPSKIIVASTTTSGEIKTSLITSGWGLA 360
      |||
Qy 361 RHFHYPELISAFETWVPALFDFNFAYFVITLTLFDRAKRDDCRSKYKWKLYC 420
      |||
Db 361 RHFHYPELISAFETWVPALFDFNFAYFVITLTLFDRAKRDDCRSKYKWKLYC 420
Qy 421 EK 422
      ||
Db 421 EK 422

```

RESULT 2

A36427
Lamin B receptor - chicken
C:Species: Gallus gallus (chicken)
C:Date: 12-Apr-1991 #sequence_revision 12-Apr-1991 #text_change 21-Jul-2000
C:Accession: A36427
R:Norman, H.J.; Evans, C.D.; Blobel, G.
J. Cell Biol. 111, 1535-1542, 1990
A:Title: The lamin B receptor of the nuclear envelope inner membrane: a polytopic protein
A:Reference number: A36427; MUID:91009487; PMID:2170422
A:Accession: A36427
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-637 <WOR>
A:Cross-references: GB:Y00822; NID:962933; PIDN:CAA68758.1; PID:962934
C:Keywords: phosphoprotein; transmembrane protein

Query Match

Best local similarity 21.4%; Score 554; DB 2; Length 637;
Matches 138; Conservative 60; Mismatches 179; Indels 40; Gaps 11;

```

Qy 18 LLAFCPPVILLMTYVHODGSVTQF-----GTFWENGVOGLINIMPRPLIAKTIIF 71
      |||
Db 210 MAFEPFATVLYLVMCKDDPSIMNFPPLPALESIMETKRGVFLM-----FF 259
Qy 72 CYGAFPAIIQLLPGKRVES-PISPAQNRPVYKANGLAIFYVTLATYLGIMWGIENPAI 130
      |||
Db 260 ----FOALFPLLPYIGKRVESPLSNP-RKIQYRNGFYAFELTAAIGTLTYFQ-PELHY 313
Qy 131 VYDHLGEIFSAIFGSIFCVLLYIKGHVAPSS--SDSGSCGNLIDFYWGMELYPRIK 188
      |||
Db 314 LYDFEVOFAVSAAFSAFSAIYLYIRSLKAEEDLAPGNGSYLVDFPTGHELPRIQ- 372
Qy 189 SFDIKVFTNCRFGMSNAVLAVTYCIKQYEINGKVSDSM---LYNTILMLVYTKFEWME 245
      |||
Db 373 SFDIKVFTNCRFGMSNAVLAVTYCIKQYEINGKVSDSM---LYNTILMLVYTKFEWME 431
Qy 246 AGYWTMTDIADHRCGFYICWGLVWVPSVYTPGMVYVNHVPELGTOLAIYILVAGILCI 305
      |||
Db 432 EAVLTMTDIHDFGFMALFAGDLVWVPYYSLOAFYLVGHPRALISMPVAAITILNCIGY 491
Qy 306 YINYDCDROQROEFRTNKGKLVWRAPSKIIVASTTTSGEIKTSLITSGWGLANRHF 365
      |||
Db 492 YIFRSANSOKNFRKN-----PADPKLSYLKLVIPATGKGLLVGMWGFVHNPY 541
Qy 366 YPELISAFETWVPALFDFNFAYFVITLTLFDRAKRDDCRSKYKWKLYC 422
      |||
Db 542 YGDLIMALWSLPCGFNHLIPFYVITYFTCLVHREARDEHCKRKYGLAMEYICOR 598

```

RESULT 3

JC5567
Lamin B receptor - rat
N:Alternate names: nuclear localization signal-binding protein, NBP60
C:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 05-Nov-1999
C:Accession: JC5567; PC4483
R:Kawahire, S.; Takeuchi, M.; Gohshi, T.; Sasagawa, S.; Shimada, M.; Abe, J. Biochem. 121, 881-889, 1997
A:Title: CDNA cloning of nuclear localization signal binding protein NBP60, a rat homolog signals and chromatin.

A:Reference number: JC5567; MUID:97335934; PMID:9192729

A:Accession: JC5567
A:Molecule type: mRNA
A:Residues: 1-620 <KAW>
A:Cross-references: DDBJ:AB002466; NID:92204061; PIDN:BAA20471.1; PID:d1021305; PID:9
A:Experimental source: Liver
A:Accession: PC4483
A:Molecule type: protein
A:Residues: 2-28,41-50,114-126,153-180,600-619 <KA2>
C:Comment: This protein is the center of perinuclear membrane network on the inner nu
C:Genetics:
A:Gene: NBP60
C:Keywords: DNA binding; nucleus; phosphoprotein; receptor; transmembrane protein
F:73-99/Region: arginine/serine-rich
F:214-235/Domain: transmembrane #status predicted <TM1>
F:256-283/Domain: transmembrane #status predicted <TM2>
F:301-332/Domain: transmembrane #status predicted <TM3>
F:331-351/Domain: transmembrane #status predicted <TM4>
F:390-411/Domain: transmembrane #status predicted <TM5>
F:415-437/Domain: transmembrane #status predicted <TM6>
F:467-488/Domain: transmembrane #status predicted <TM7>
F:557-587/Domain: transmembrane #status predicted <TM8>
F:45/Binding site: phosphate (Thr) (covalent) (by calmodulin-dependent kinase II) #st
F:71/Binding site: phosphate (Ser) (covalent) (by cdc2 kinase) #status predicted
F:78/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted
F:123/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status predicted
F:137/Binding site: phosphate (Tyr) (covalent) #status predicted
F:161/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase II) #s
F:206/Binding site: phosphate (Thr) (covalent) (by cdc2 kinase) #status predicted

Query Match

Best local similarity 21.3%; Score 553.5; DB 2; Length 620;
Matches 138; Conservative 65; Mismatches 176; Indels 39; Gaps 11;

```

Qy 16 LSLAFCPPVILLMTYVHODGSVTQ-----TFGFWENGVOGLINIMPRPLIAKTI 69
      |||
Db 221 LLIMGLPACVFLLLLOCAKDPGLLOFPPLPALREIMEARVCGVILM----- 270
Qy 70 IFCYGAFAIIQLLPGKRVES-PISPAQNRPVYKANGLAIFYVTLATYLGIMWGIENPA 129
      |||
Db 271 FF----LOALFSLPLVGVWVEGTPYVDBGRRLKYLKNGLYAFILTSAA-VGTAYFWIDELY 325
Qy 130 IYDHLGEIFSAIFGSIFCVLLYIKGHVAPSSDS--GSCGNLIDFYWGMELYPRIK 188
      |||
Db 326 YLYTHFQLFALAAIYFVSVLYARSLKYRDELSPASSGNAVYDFIGRELNPRIQ- 384
Qy 189 SFDIKVFTNCRFGMSNAVLAVTYCIKQYEINGKVSDSM---LYNTILMLVYTKFEWME 245
      |||
Db 385 AFDKFPELPRGLIGWVIMVLMLEMKVQERSASPSLMTLVNS--FOLLVVDALMFE 443
Qy 246 AGYWTMTDIADHRCGFYICWGLVWVPSVYTPGMVYVNHVPELGTOLAIYILVAGILCI 305
      |||
Db 444 EAVLTMTDIHDFGFMALFAGDLVWVPYYSLOAFYLVGHPRALISMPVAAITILNCIGY 502
Qy 306 YINYDCDROQROEFRTNKGKLVWRAPSKIIVASTTTSGEIKTSLITSGWGLANRHF 364
      |||
Db 503 YIFRCANSOKNAFRKN-----FTDKLHLNLTIPSTYKSLSVSGMWGFVHNPY 552
Qy 365 YPELISAFETWVPALFDFNFAYFVITLTLFDRAKRDDCRSKYKWKLYC 422
      |||
Db 553 YGDLIMALWSLPCGFNHLIPFYVITYFTALLIHREARDEHCKRKYGLAMEYICOR 610

```

RESULT 4

A53616
Lamin B receptor - human
N:Alternate names: LBR
C:Species: Homo sapiens (man)
C:Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 05-Nov-1999
C:Accession: A53616; A53617
R:Ye, Q.; Norman, H.J.
J. Biol. Chem. 269, 11306-11311, 1994
A:Title: Primary structure analysis and lamin B and DNA binding of human LBR, an inte

us-09-817-774-31.rpt

[illegible]

Rilvine, M.; Wood, V.; Rajadream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.
Submitted to the EMBL Data Library, May 1998
A:Reference number: Z1866
A:Accession: r73610
A>Status: preliminary; translated from GB/EMBL/DDbJ
A:Molecule type: DNA
A:Residues: 1-424 <LN>
A:Cross-references: EMBL:AL023554; PIDN:CAI19037.1; GSPDB:GN00067; SPDB:SPBC165.18
A:Experimental source: strain 972h; cosmid c1655
C:Genetics:
A:Gene: erg24
A:Map position: 2
C:Superfamily: yeast probable C-14 sterol reductase
C:Keywords: oxidoreductase

Query Match 19.3%; Score 500.5; DB 2; Length 424;
Best Local Similarity 32.5%; Pred. No. 1.le-34;
Matches 128; Conservative % 67; Mismatches 172; Indels 27; Gaps 11;

OY 43 TGFGEWENGVOG-----LINIMPRPTLIAMKIIIFCYGAFEALLOL---LPGRVEG- 91
|||:::||::||::||::||::||::||::||::||::||::||::||::||
Db 32 SFGSYITONEECCKPAKTSKISHIKTKIPLEDQSLIYLLEFSTLLMLKCTGNKAAGT 91
OY 92 PISPAGNPVRYKANGL-AAEVLATATLGIMWFSEFNDAIVDH-LGEIRSALESFTF 149
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 92 PLDDGTGLTXKINGFNACILIGVCTSIYLLGASCMERFMNFIQLMAAAVF-SVAL 150
OY 150 CVLLXIKKHVA PSS-SDSGSGCN.IIDPYMGELPPIGSDFKVTNCRFGMSNAVU 208
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 151 CFPCVGSFFSKQOLAKGTSIGNILFDWFIGRSLNPIRG-NFDICFCCLRPGLIMVAF 209
OY 209 AVTYCIKKY-EINGKVS DSM.LVNITLM.LVYTTFEFMEAGVMTMDIAHNRGFCYICWG 267
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 210 DIAFHCHQELVLRGRITDSWLVIIHFHTWYVDSLINESAVLTMDITDDEGYMLSFED 269
OY 268 LVMWPVSVTSGMFLVNHPELGTOLAIIYLIVAGLICITYNYDCDROROEFRFTNGKCLV 327
|||||::|||||::|||||::|||||::|||||::|||||::|||||::|||||
Db 270 LVMPVELISLARIALAHFPVDLGLVKTIALLCLQFLFYIFRGCANCGKNFRSNPNM-- 326
OY 328 WGRAPSKIVASVTTSGTEKTSLTLTGSGWMGARHHEVPDELISAPFWVPALDFNFLAY 387
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 327 ----PKLKLHFLTQTKRKTK----LLISGMWKNAHNIVFGDMIMAWMACPLPAGFSPIPY 379
OY 388 FYVIFLTLLFDRAKRDDRCRSKYGYMKLYCE 421
|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 380 FYVAFGVLLVHRNARDDHKCRVXYGEDWEKYCK 413

RESULT 6
S44170
Probable C-14 sterol reductase (EC 1.1.-.-) - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #extl_change 17-Mar-2000
C:Accession: S44170
submitted to the EMBL Data Library, March 1994
A:Reference number: S44170
A:Accession: S44170
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-490 <PAP>
A:Cross-references: EMBL:X77955; NID:g473245; PID:g473246
C:Genetics:
A:Introns: 26/2

Query Match 18.3%; Score 473.5; DB 2; Length 490;
Best Local Similarity 30.9%; Pred. No. 2.5e-32;
Matches 133; Conservative 61; Mismatches 151; Indels 85; Gaps 18;

48 W-ENGVGGGLINIMPRPLIAMKIIFCYGAFEALLOLPGKRVREGSPIAGNRNPYKANG 106
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 75 WPDNGVFGFLS-W-----SATLWTL--GYVALSVQYRFLPGHVEGTELSTGGRKLKINA 128

QY 107 LAAYEYTLATYIG-----LMMFGLFNPALIVYDHLGELFSALLIGSFIFCYLLYIK 156

Db 129 FNSAMCTLALIAAGTIAOGEFPPVTF-----ISDNFQIISANILFAFALAFVYVR 181

QY 157 G-HVAPSSD-----SGSGCNLIIDFYGMELYPRIQ---KSPDIFVTNCRGMSMW 205

Db 182 SDYKFGKNDKMQALAGVYTGSLITDFYIGRELNRITPLIGQVDIKEFPMRGLLGW 241

QY 206 AVLAVTYCIKOYEINGKVSMSLVNTILMLVYTRFFWMEAGYWNMTMDIAHDSGFYICW 265

Db 242 IILNCAFIKQYRLXYVDSITLITAIQAFYFEDGIWEPALVLTMDITTDGFGFMSLF 301

QY 266 GCLVWPSPYITSPGMLVNHPELGT--QLAI-YIIVAGILCIYINQDROROFRRFN 322

Db 302 GDVWVWFYFSTQTRFLSVHPQOLGAFGLIAGVLAAGYSIFRLS---NSQKNF-RFN 357

QY 323 GKCLVWGRAPSKIVASYTTTSGEKT-SLLTSGWMLGRHHPYVELLSAFWTVP--- 378

Db 358 PE-----DPYKHLTYL-----QTKGSRLLITSGWMLGRHHPYVELLSAFWTVP--- 407

QY 379 -----ALFDN-----FLATYVIFLULLFDRAKRDDRCR 409

Db 408 AGYQILSAGSNAPGATIMLDGREVVQGEARGWGIYFTYFYLIFAILLIRDLRDERKS 467

QY 410 SKYGYWKLY 419

Db 468 KKYGDWERY 477

RESULT 7

S30769

Probable C-14 sterol reductase (EC 1.1.-.-) - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein N0593; protein YNL280C

C:Species: *Saccharomyces cerevisiae*

C:Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 20-Jun-2000

C:Accession: S30769; S63254

R:Jorenz, R. T.

submitted to the EMBL Data Library, November 1992

A:Description: Cloning, sequencing, and disruption of the gene encoding sterol C-14 red

A:Reference number: S30769

A:Accession: S30769

A:Molecule type: DNA

A:Residues: 1-438 <LOR>

A:Cross-references: EMBL:M99419; NID:g171142; PID:g171143

K:Jessen, F.; Dubois, E.; Vierendeels, F.; Scherens, B.; Pierard, A.; Glansdorff, N.

submitted to the Protein Sequence Database, April 1996

A:Reference number: S63254

A:Accession: S63254

A:Molecule type: DNA

A:Residues: 1-438 <MES>

A:Cross-references: EMBL:271556; NID:g1302350; PID:g1302351; MIPS:YNL280C

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:ERG24

A:Cross-references: SGD:S0005224; MIPS:YNL280C

A:Map position: 14L

C:Superfamily: yeast probable C-14 sterol reductase

C:Keywords: oxidoreductase; transmembrane protein

F:17-33/Domain: transmembrane #status predicted <TM1>

F:74-90/Domain: transmembrane #status predicted <TM2>

F:109-125/Domain: transmembrane #status predicted <TM3>

F:145-161/Domain: transmembrane #status predicted <TM4>

F:247-263/Domain: transmembrane #status predicted <TM5>

F:312-328/Domain: transmembrane #status predicted <TM6>

Query Match 16.4%; Score 424.5; DB 2; Length 438;

Best local similarity 29.7%; Pred. No. 3.2e-28;

Matches 130; Conservative 74; Mismatches 160; Indels 73; Gaps 20;

Db 28 PVFTIIL-NQMRPDIYK---GPFQNFIDVLEW-NGIKPLRYLGNREL----- 73

QY 69 IIFC--YGAFAIILQLLPGRKVBGPISPAKNRPYKANKLAVFYTLATYIGLMMFGIF 126

Db 74 TVYGLWGLL-AVLDVILPGRVWKMGVQLRDSKLSYKINQI-AMSTVLVLAIRKRLTD 131

QY 127 N-PAIY-----DHGEIFSAIIFGSFI-----FCVLLYIK---GHVAPSSSDSGSG 170

Db 132 GQPELDQYVENHVSCLITSLIFSPFLATYCYASFPILFKKNGKRRKILALGNSG 191

QY 171 NLIDFYGMELYPRIKSPDIDKVTNCRGMSMAVLAITYCIKOYEINGKVSMSLVN 230

Db 192 NIIDWETGRELNRILG-PIDIKMFSELRGMLIMLILINISCHHHNLTGKINDALVLY 250

QY 231 TIILVYVTRFFWMEAGYWNMTMDIAHDSGFYICWGLVWVSPYITSPGMLVNHPELG 290

Db 251 NFDGFIYFDGVLEEBEVLNMTDITTDGFGFMAFGLSLVPFTYISQARTLSVPELG 310

QY 291 TOLAVIYIVAGILCI-----YINYCDROROFRRNKGKLVWGRAPSKIVASYTTSGE 345

Db 311 -----WVKVVGILAIIMFLGFHIFHSANKQSEFRQ--GKL-----ENLKSIGTRKG- 354

QY 346 TKTSLLTSGWMLGRHHPYVELLSAFWTVPALFNPFLATYVIFLULLFDRAKRDD 405

Db 355 TK---LLCDGWMARSOHINFGDWLISLWCLATWTFQTPLYYSIXFATLLHROORDE 411

QY 406 DCRSKYGYWKLYCEK 422

Db 412 HKCRKTYGEMWEYERK 428

RESULT 8

T18681

hypothetical protein B0250.9 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T18681

R:Lloyd, C.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19005

A:Accession: T18681

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-697 <WIL>

A:Cross-references: EMBL:Z81453; PIDN:CAB03797.1; GSPDB:GN00023; CESP:B0250.9

A:Experimental source: clone B0250

C:Genetics:

A:Gene: CESP:B0250.9

A:Map position: 5

A:Introns: 211/3; 281/3; 318/3; 372/1; 436/1; 490/1; 618/1

Query Match 16.1%; Score 418.5; DB 2; Length 697;

Best local similarity 29.3%; Pred. No. 1.7e-27;

Matches 122; Conservative 76; Mismatches 200; Indels 19; Gaps 9;

QY 10 VTYASMSLAFCEPPEVILMYTVNHODG---SVTQPFGEWENGVOGLINIMPRPTLI 65

Db 286 VSAQWVALLIIVPPAPFLEFYSISIHGTGLVPTIFALFLRF-PLVLCQVPPVW---DTV 341

QY 66 AMKIIFCGAFEDAILQLLPKRVREGPISPAKNRPYKANKLAVFYTLATYIGLMMFGI 125

Db 342 AKWESAVNCAIQLLFYVWLPDQAL-VWSSAGDQ-MREVNSEFSCILITCLLYVIGASAGV 399

QY 126 FNPALVYDHLGEISALIFGSFICVLLYIGHVAAPSDDSGSCGNLIIDFYWMGELYP 185

Db 400 YRGDLVYHNSIIT--LIFA--IFAVLIW-----AALATYHFGVITYIISSEWFGIDENHPK 451

QY 186 IGSFIDIKVTNCRGMSMAVLAITYCIKOYEINGKVSMSLVNTIIMLVYVYKFFWE 245

Db 452 I-LDIDKASFIRTRFTYIWPFLVISAMYFKKITYGQISVLCVSSVOLLTYIQFPMHE 510

QY 246 AGYNTMDIAHDSGFYICWGLVWVSPYITSPGMLVNHPELGTOLATYIIVAGILCI 305

Db 511 DLFNLSKRCDFGFIYIMADFNVLGPIITYSPTIIVAINRSGVISNCLFCAVAIGSM 570
 QY 306 YINYDCDRQROEFRTNCKLWGRAPSKIYASTTTSGTSTLLTSGWGLARHPHY 365
 Db 571 VETACDQKQKYEFRKSTLAVGVDAFFISAKYRTDSGDANTLLSGHWGCKRHNY 630
 QY 366 VPEILSAFFWVPALFDNFALFYVIFLTLLEFRAKRDDRCRSKYKWKLYCEK 422
 Db 631 ASEAITFAAFSAFOGFSTIAHPSILFYFLFLVARAFDENRCLIKKGQVWQYCSK 687

RESULT 9

147551
 nuclear envelope membrane protein-like - Arabidopsis thaliana
 N:Alternate names: protein F8J2.110
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000
 C:Accession: T47551
 R:Nyakatura, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.; Mayer, K.F.X.
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: 224458
 A:Accession: T47551
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-317 <N>A
 A:Cross-references: EMBL:AL132969
 A:Experimental source: cultivar Columbia; BAC clone F8J2
 C:Genetics:
 A:Map position: 3
 A:Insertions: 12/3; 52/1; 75/3; 94/3; 125/2; 141/2; 176/3; 202/2; 242/2; 267/2
 A:Note: F8J2.110

Query Match 14.3%; Score 371.5; DB 2; Length 317;
 Best local similarity 28.7%; Pred. No. 6.9e-24;
 Matches 107; Conservative 68; Mismatches 107; Indels 91; Gaps 16;

QY 62 PTLIAKIIIFCYGAFAIQLLLPGKRVGEPISPAKRPYKANGLAAYFVTLATYLGW 121
 Db 8 PSLGSVVLVEFYVYLAAGELIPGKVIKRGVLLSDGSLRRCGLLA-LILVALILGIC 66
 QY 122 -WEIFNPALVYDHLGEISALIGSFICYL--LYIKGHVAPSSDSGS-----CG 170
 Db 67 AKLIVSPVLVADRGLELSA---TFICVLTLLIYVGR---SSSNKSGSLKPHVSG 119
 QY 171 NLIDFWMGMLYPRICKSPDIKVFYTCNRFQMGSMNAVLAATYCTCKQKYEINXVSDSLVN 230
 Db 120 NLVHDWFGIQLNPOF-MSIDLKA-----GMMGMLINLSILAKSVQ-DGSLQSOMILY 171
 QY 231 TILMLVYVTKFEFWMAGYVWMTIDAHDRGFIYICWGLVWVPSVYTSFGMYLVNHPYELG 230
 Db 172 QIFC-----AGMW-----LLANKVE-- 166
 QY 291 TQLAIYLVAGILCIYIN-----DCDRQROEFRTNCKLWGRAPSKIYASTTTSGET 346
 Db 187 --LIVPAIVNCVLEFLGYVWFRGANKOKHFKK-NKPTIWGKRP-----VVVGGK- 235
 QY 347 KTSLLTSGWGLARHPHYVEILSAFFWVPALFDNFALFYVIFLTLLEFRAKRDD 406
 Db 236 ----LLVSGYVIGIARHCNYIGADMLALSLFPCGISSPVFYFYLILLIMRERDEV 291
 QY 407 RCRSKYGRWKLY 419
 Db 292 RCAEKYEIMAEY 304

RESULT 10

S64014
 probable transpore protein ERG4 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein G3725; hypothetical protein YGL012w; hypothetical
 C:Species: Saccharomyces cerevisiae
 C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 21-Jul-2000
 C:Accession: S64014; S15042; S31559

R:Hebling, U.; Hofmann, B.; Delius, H.
 submitted to the Protein Sequence Database, May 1996
 A:Reference number: S64003
 A:Accession: S64014
 A:Molecule type: DNA
 A:Residues: 1-473 <HEB>
 A:Cross-references: EMBL:472534; NID:g1322468; PID:g1322469; MIPS:YGL012w
 A:Experimental source: strain S288C
 R:Chen, W.; Capleau, E.; Balzi, E.; Goffeau, A.
 Yeast 7, 305-308, 1991

A:Title: The YGL022 gene encodes a putative transport protein.
 A:Reference number: S15042; MUID:91353085; PMID:1882555
 A:Accession: S15042
 A:Molecule type: DNA
 A:Residues: 1-365, 'V', 367-473 <YEA>
 A:Cross-references: GB:S57891; NID:9234324; PID:AA19615.1; PID:9234325
 A:Experimental source: strain IL125-2B
 R:Chen, W.; Balzi, E.; Capleau, E.; Choder, M.; Goffeau, A.
 Yeast 7, 287-299, 1991
 A:Title: The DNA sequencing of the 17 kb HindIII fragment spanning the LEU1 and ATE1
 ntrolling plasmid drug resistance.

A:Reference number: S15040; MUID:91353083; PMID:1882553
 A:Accession: S31559
 A:Status: translation not shown
 A:Molecule type: DNA
 A:Residues: 1-365, 'V', 367-473 <CHE>
 A:Cross-references: GB:S58126; NID:9234321; PID:AA13895.1; PID:94261595
 C:Genetics:
 A:gene: SGD:ERG4
 A:Cross-references: SGD:S0002980; MIPS:YGL012w
 A:Map position: 7L
 C:Superfamily: yeast probable C-14 sterol reductase

C:Keywords: transmembrane protein
 F:46-62/Domain: transmembrane #status predicted <TM1>
 F:103-119/Domain: transmembrane #status predicted <TM2>
 F:153-169/Domain: transmembrane #status predicted <TM3>
 F:175-191/Domain: transmembrane #status predicted <TM4>
 F:330-346/Domain: transmembrane #status predicted <TM5>
 F:424-440/Domain: transmembrane #status predicted <TM6>

Query Match 14.3%; Score 371; DB 2; Length 473;
 Best local similarity 26.4%; Pred. No. 1.2e-23;
 Matches 115; Conservative 71; Mismatches 202; Indels 48; Gaps 15;

QY 15 MSLAFCPPVLLMYTMVHODGSYT-OTFGFW-----ENGVOGLINIMPR 61
 Db 46 VIGMLIGFPLMLYMWICAEFYHGKVALPKAGESMWHFKHLYOVLNENGI----- 96
 QY 62 PTLIAKIIIFCYGAFAIQLLLPGKRVG-PISP-AGNRPYKANGLAAYFVTLATYLG 119
 Db 97 PEKYDWTIFLTFWVPOIIFYYITGPGTWGQPLSHKQQLYFCNAMTLYVTTTLVLV 156
 QY 120 LWMFGLFNPALVYDHLGEISALIGSFICYLVLYIKGHVAPSSDSGSGNLIIDFYWG 179
 Db 157 LHFTNLFRLVLIIDRGRIMTCAIISGFASIIYLV-WTLFISHDYHRMTGHNHLDFFMG 215
 QY 180 MELYPRICKSPDIKVFYTCNRFQMGSMNAVLAATYCTCKQKYEINXVSDSLVNITILMYVT 239
 Db 216 APLNPRWG-ILDKMEFEVRLPWFYLTTLTGACLKQWETGYVYVPLQGLVWLAHMLVN 274
 QY 240 KFEFWMAGYVWMTIDAHDRGFIYICWGLVWVPSVYTSFGMYLVNHPYELG----TOLA 294
 Db 275 ACAKGEEILVPTMDAYEKGFMLIFMNLAGVPIYCHCTILVLYHDPSEYIMSTLYNVS 334
 QY 295 IYLVAGILCIYINND-CDROEER-----RNGKCLWGRAPSKIYAS--YTTSGE 345
 Db 335 LYVV--LLCAYYFFDTANAOKNAFRKQMSGDKTGKTPF--LPYQLIKNKYVWTS-- 387
 QY 346 KTSLLTSGWGLARHPHYVEILSAFFWVPALFDNFALFYVIFLTLLEFRAKRDD 405
 Db 388 -NGSYLLIDGMVTLTARKHITADWQSLWALSCGNSVFFVFFVFLVYLHRAFDQ 446
 QY 406 DRCRSKYGRWKLYCE 421

Db 447 AKCKRKYKMDYCK 462

||:|||||:|:

RESULT 11

A43765

stai1 protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 18-Nov-1992 #sequence_revision 18-Nov-1992 #text_change 01-Dec-2000

C:Accession: A43765; J38121

R:Shimanuki, M.; Goebel, M.; Yanagida, M.; Toda, T.

Mol. Biol. Cell 3, 263-273, 1992

A:Title: Fission yeast stai1(+) gene encodes a protein similar to the chicken lamin B

A:Reference number: A43765; MUID:92329994; PMID:1320960

A:Accession: A43765

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-453 <SH1>

A:Cross-references: GB:X63549; NID:95108; PIDN:CAA45113.1; PID:95109

R:Connor, R.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1996

A:Reference number: Z21771

A:Accession: T38121

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-411, 'C', 413-453 <CON>

A:Cross-references: EMBL:Z98600; PIDN:CAH11256.1; GSPDB:GN00066; SPDB:SPAC20G4.07C

A:Experimental source: strain 972h; cosmid c20G4

C:Genetics:

A:Gene: SPDB:SPAC20G4.07C

A:Map position: 1

C:Superfamily: yeast probable C-14 sterol reductase

C:Keywords: transmembrane protein

Query Match 14.0%; Score 363.5; DB 2; Length 453;

Best Local Similarity 26.8%; Pred. No. 4.9e-23;

Matches 117; Conservative 69; Mismatches 204; Indels 47; Gaps 12;

QY 16 LSLAFCPPVILLMTVMVHODGSVOTGCF--WENGVOGL--INIMPRPTLAWKI 69
 Db 20 LAIMTGPCLMYLMSKRNDSQFTKRESFTAGQNFRTIGHYIYGAVPTRAFLV 79
 QY 70 IFCYGAFEALQLLPGKRVES-PISPAGN-RPVYKANGLAAYVTLATYLGIMFGIFN 127
 Db 80 FMSFCIAQAVMYLTLPGVFTQGLPKHRNNEBRLPYLCNMIWSFTYIVLAVLHTVHP 139
 QY 128 PAIVYDHLGEIFSAIFGSIIFCVLXI-----KGHAPSSSDSGSGNLIIDF 176
 Db 140 ITTFIDMFGPLMSVAITTAFTVCTFVLYTGLTFGRLLPDKPH-----RLSGNPIYDA 191
 QY 177 YMGELAPRIGKSFEDIKVFNCRFGMMNAVL--AVTICIKOYEINGKVSMSLVNTIL 233
 Db 192 FMGACLNPRIGKLLDFKMFVFR---IPWILFISVGAANRKYETGYVSPVLFVCG 248
 QY 234 MLVYVTFEWMWAGYNTMDIAHNRGEFYICWGLVAVDSVYTPGMVYLVNHEVEL--G 290
 Db 249 HYLVAANCKGEOLIVPTMDMAYEKFGFMLIFWNMAGVFPYTSCHTLVLFSDPSYNNMS 308
 QY 291 TOLAIVLVAGLICITVNYCCDROEFR---TNGKCLVWRASKIVAST--TTS 343
 Db 309 TQYTTGIVYLLCCYYIEDTCNGCKNHFNOIYGTVEHRTKFPQDLWLIKNEPTFCAN 368
 QY 344 GEFTSLTLTSGWMLGARHPHYVEILSAFWTPALFDNFALYFVIFETLLDFRAR 403
 Db 369 GGI-----LLTSGMYRARKIHTYADFPQSLSMALITGFQSPLEFYFSPFFVVLVHVR 424
 QY 404 DDDCRSKYKWKYC 420
 Db 425 DIKKCKAKYGADEYCK 441

RESULT 12
S14138

cytochrome-c oxidase (EC 1.9.3.1) chain I - sugar beet mitochondrion
 C:Species: mitochondrion Beta vulgaris var. altissima (sugar beet)
 C>Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 23-Jul-1999
 C:Accession: S14138
 R:Sendra, M.; Harada, T.; Mikami, T.; Sugiyama, M.; Kinoshita, T.
 Curr. Genet. 19, 175-181, 1991
 A:Title: Genomic organization and sequence analysis of the cytochrome oxidase subunit
 A:Reference number: S14138; MUID:91330331; PMID:1651175
 A:Accession: S14138
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-524 <SEN>
 A:Cross-references: GB:X57693; GB:S47702; NID:911258; PIDN:CAA40874.1; PID:911259
 C:Genetics:
 A:Gene: coxi
 A:Genome: mitochondrion
 C:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C:Keywords: chromoprotein; copper; electron transfer; heme; iron; magnesium; membrane
 F:12-459/Domain: cytochrome-c oxidase chain I homology <CO1>
 F:64-380/Binding site: heme a iron (His) (axial ligands) #status predicted
 F:243-292,293/Binding site: heme a iron (His) #status predicted
 F:243-247/Cross-link: 1'-histidyl-3'-lysine (His-Tyr) #status predicted
 F:247/Binding site: oxygen (Tyr) #status predicted
 F:370/Binding site: magnesium (His) (shared with chain II) #status predicted
 F:378/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 4.6%; Score 120.5; DB 2; Length 524;

Best Local Similarity 20.4%; Pred. No. 0.022;

Matches 110; Conservative 56; Mismatches 169; Indels 205; Gaps 28;

QY 64 LIAW-----KIIFCYGAFAEI-----LQLLPGKRVAGSPISPAGNRPVYK 103
 Db 4 LVRMLFSTNHMDIGTYIRFPAIAGVGTCSVLIRBELAHPGQIIG-----GNHQLVN 58
 QY 104 A-----NGLAIEYATLATYGLM--WFG---TFNPAIYDHLGEIFSAIFGSIIFCVL 153
 Db 59 VLITAHAFIMTFEWMWAGYNTMDIAHNRGEFYICWGLVAVDSVYTPGMVYLVNHEVEL--G 114
 QY 154 YIKGHVAPSSSDSGSGNLIIDFYWMGMEIYPR-----GKSPDIKVF----- 195
 Db 115 LLSALVYVSGSGTG-----W--TVYPLSGITSHSGAVDLAIFSLHLSGVSSIL 162
 QY 196 -----TNCRFGMMAVAIATVTCIKOYEINGKVSMSLVNTILMYV 238
 Db 163 GSINFTITTFNMRGPGMHRRLPFVMSVLTAFEL-----LLSLPVLATITMLTFD 215
 QY 239 TKFWMWAGYNTMDIAHNRGEFYICWGLVAVDSVYTPGMVYLVNHEVEL--G 288
 Db 216 RNF-----NTTFEPDAGGDDPLLYOHLFWFGHPEVYLLLPGRGILSHIVSTESGK 267
 QY 289 -----LGTOLA-IYTLVAGIL--CIYNYDDROEFR-----TNGKCLVWG 329
 Db 268 PVEGYLCAVYAMISIGVIGFLVMAHMFVGLDAD--TRAYTAATMILAVPTGKIISW- 325
 QY 330 RAPSRIVASTYTTSGETKTSLLTSGW-----WGLARHF 363
 Db 326 -----IATMGGSIQYKTPMLFAVGFIPLFTVGIGITIVLANGDLIALHDYVVAHF 379
 QY 364 HYV-----PELISAF-FWTVALDNDNLAAYVYFETLL 397
 Db 380 HYLISMGAVALFAGFYVWGIKFGRIYPERLGOIHFI--TFEGVNLTFPPMHFLGSG 437
 QY 398 FDRARDDDRCKSKYKWKLYCEKSNSTGSPREFIDCEWCCSHFLTLTLESSKDR 457
 Db 438 MPRRIIPD-----YPDAYAGNMLSSGYSISVIGICF-FYVYVIT--TSSGKNK 484

RESULT 13
S65346
cytochrome-c oxidase (EC 1.9.3.1) chain I - tomato mitochondrion
 C:Species: mitochondrion Lycopersicon esculentum (tomato)
 C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000

Page 7

QY 154 YIKGVAPSSSDSGSCGNLLIDFTWGMELYPRI-----GKSEDIKVF-----195
 Db 115 LLSALVEVSGGTG-----W--TVPPLSGITSHSGAVDALIFSLHSGVSSIL 162
 QY 156 -----TNCRCGMMSMVAIVATYCIKQYEINGKXSDSMVNTIIMLYV 238
 Db 163 GSINITITTFNNRCRGMTMHRLPFWSVLYTAFL-----LISLVLACALIMLLTD 215
 QY 239 TKFFWMEAGYNTMDIADRGGFYICWGLVWV---PSVY--TSPGMVNNHPE-----288
 Db 216 RNF-----NTTFPDAGGDDPLLYQHLEFPEBHPREYILILPFGIISHIVTSESG 267
 QY 289 -----LGTOLA-TYLLVAGIL-----CIYINDCDROBOEFRR-----TNGKCLVWG 329
 Db 268 PVFGYLGWVYAMISIGVLGFLWAMHMEFTVLGDV-TRAYFTAMTIMIAYPLGIKFSM- 325
 QY 330 RAPSKIVASYTTSGEFTSLITLSGM-----WCLAHNF 363
 Db 326 -----IATMMGGSIQYKTPMLFANGFIFLFTIGGLTGIVANSGLDIALHDYIVVAHF 379
 QY 364 HYV-----PELISAF-FWTVPALPDEFNLAYEVYIFLTL 397
 Db 380 HYVLSMGAVFLAFAGFYWVGKIFGRTYPELLGQHNFI--TFEGVNMTFFPMHGLGSG 437
 QY 398 FDRARDDRCRSKTKGKWKVLCXKSNSTGSRREFIDCEVCCSHFLTILRTLESSDR 457
 Db 438 MPRIIPD-----YPDAYAGMNAISFGSYISVVGICG-FEVVVTI--TLSSGRNK 484

RESULT 14
 S14139
 cytochrome-c oxidase (EC 1.9.3.1) chain I - radish mitochondrion
 C:Species: mitochondrion Raphanus sativus (radish)
 C:Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 23-Jul-1999
 C:Accession: S14139
 R:Makarov, C.A.; Apel, I.U.; Palmer, J.D.
 Curr. Genet. 19, 183-190, 1991
 A:Title: The role of coxi-associated repeated sequences in plant mitochondrial DNA re
 A:Reference number: S14139; MUID:9130332; PMID:1651176
 A:Accession: S14139
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-527 <MAK>
 A:Cross-references: GB:X57692; NID:g297415; PIDN:CAA40873.1; PID:g297416
 C:Genetics:
 A:Gene: coxi
 A:Superfamily: cytochrome-c oxidase chain I; cytochrome-c oxidase chain I homology
 C:Keywords: chromoprotein; copper; electron transfer; heme; iron; magnesium; membrane
 transmembrane protein
 F:12-459/Domain: cytochrome-c oxidase chain I homology <COL>
 F:64,380/Binding site: heme a iron (His) (axial ligands) #status predicted
 F:243,292,293/Binding site: copper (His) #status predicted
 F:243-247/Cross-link: 1--histidyl-3--tyrosine (His-Tyr) #status predicted
 F:247/Binding site: oxygen (Tyr) #status predicted
 F:370/Binding site: magnesium (His) (shared with chain II) #status predicted
 F:378/Binding site: heme a3 iron (His) (axial ligand) #status predicted

Query Match 4.58; Score 115.5; DB 2; Length 527;
 Best Local Similarity 20.2%; Pred. No. 0.058;
 Matches 109; Conservative 56; Mismatches 170; Indels 205; Gaps 28;

QY 64 LIAM-----KIIFCVGAFAEI-----IQLLLPGRKRVGCPISPAONRRVYK 103
 Db 4 LVRLPFSNNKDDITGLTFIFGALINGWGTGCSVLRRELAPRGQILIG-----GNHOLYN 58
 QY 104 A-----NGLAYEVTLATYVGLW--WEG--IFNPALVYHDLGIFSFALIFGSEIFCVLL 153
 Db 59 VLTIAHAFLMIFEMFVWPMAMGCGFNWFVPLILGNPDMAFPLNNISFWMLPPS---LLL 114
 QY 154 YIKGVAPSSSDSGSCGNLLIDFTWGMELYPRI-----GKSPDIYVF-----195

162 .

```

QY 196 -----TNCRFGMASMAVLAVTYCIKQEINGKVSQDSMLVNTILMLVYV 238
Db 163 GSINFTTTTFNNRGRPGMTNHRLLPFWASVLYVAFL-----LISLPLACATMLLTD 215
QY 239 TRFWMWAGYWNMTMDIAHNRGFIYWGCLVWY---PSVY--TSPGMYLVNHPV----- 288
Db 216 RNF-----NTFFPDPAAGGDPILYQHLFWFGHPREYIILPQFGIISHIVTFSGK 267
QY 289 -----LGTOLA-IYLYVAIIL-----CIYINDCDROREFR-----TNGCLVWG 329
Db 268 PVFGYLGMYVYAMISIGVLGFLWVAMHMTVGLVDV--TRAYFAAAMTIAVPGIIFGSM- 325
QY 330 RAPSRTVASTTTSGETKSLTSGM-----WGLARHF 363
Db 326 -----IATWGGSIQYKTPMLFANGFIPLFTIGITGIVLANSGLDIAHDTYVVAHF 379
QY 364 HVY-----PELISAF-FWTVPALFDFNLAIFYVIELTLL 397
Db 380 HYLSMGAVFALEAGFYVWVGKIFGRTPETLQIHFWI--TEFGYNLFEFFMHLGLSG 437
QY 398 FDRAKDDDRCKSKYKYLKCEKSNSTGFRFIDCNVCCSHLITLITLSSKDR 457
Db 438 MPERRIPD-----YPDAYAGWMAIASSFGSYISVWGIC--FPVYVYI--TLSSGNK 484

```

RESULT 15

S67084 probable membrane protein YOR192c - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein O4759

C:Species: Saccharomyces cerevisiae

C>Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 19-Apr-2002

C:Accession: S67084

R: Hughes, B.; Pohl, T.M.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66685

A:Accession: S67084

A:Molecule type: DNA

A:Residues: 1-599 <HUG>

A:Cross-references: EMBL:Z75100; NID:q1420458; PID:q1420459; GSPDB:GN00015; MIPS:YOR192C

A:Experimental source: strain S288C

C:Genetics:

A:Gene: MIPS:YOR192C

A:Cross-references: SGD:S0005718

A:Map position: 15R

C:Superfamily: uracil transport protein

C:Keywords: transmembrane protein

F:78-94/Domain: transmembrane #status predicted <TM1>

F:116-132/Domain: transmembrane #status predicted <TM2>

F:174-190/Domain: transmembrane #status predicted <TM3>

F:199-215/Domain: transmembrane #status predicted <TM4>

F:280-296/Domain: transmembrane #status predicted <TM5>

F:398-414/Domain: transmembrane #status predicted <TM6>

F:447-463/Domain: transmembrane #status predicted <TM7>

F:485-501/Domain: transmembrane #status predicted <TM8>

Query Match 4.2%; Score 108; DB 2; Length 599;

Best Local Similarity 18.4%; Pred. No. 0.29;

Matches 86; Conservative 172; Mismatches 172; Indels 140; Gaps 24;

```

QY 9 IYVYASMLSLAFCPPEVILLM-YTMVHODGSVTO-----TEGFW 48
Db 131 IYVYGSNAWLGGLCVNNILDSMHHLLPNTLSSKVAAMTKELGFIIFHILTAFCYFM 190
QY 49 ENGVOGLINIWPRPTLIAMKIIFCYGAFAIIOI---LLGKRVEGPISDPAGNRPYKKN 105
Db 191 KPYHMYIILWS-----CVGTFEALMGVIYILTKSAHGVODLFTSTHSTVTGSK 239
QY 106 GLAAYFTLTATYGLWMPGIFNPAIV---YDHLGE-----IFSALIFGSFICVLLIYIG 157
Db 240 KAMAWYTTIS---YWGSVSPGCTINQSDSFRGSSNCAIWTGTVALLIPATLIPVFG 294
QY 158 HVAPSSSDSGSCNLLIDFYW-GMEIYPRIGKSFIDIKVFNCRFGMMASVAVLAVTYCIKQ 216

```

```

Db 295 IIGAS-----ACEKLYGQTFWMPMDI-----EDNMLTTNYSAGARA-ATFECCGCFW 341
QY 217 YEINGKVSQDSMLVNTILMLVYVTKFEFWMEAGYNNNTDIAH-----DRGIFY---ICM 265
Db 342 SQISYTTISN-----CGFASGMDLAGLLPKRYDIKRGALFAACVSW 381
QY 266 GCLVWVPSVYTSFGMYLVNHPVELGTOLAIYIYAGILCIYINYDCR---ORQEFRRN 322
Db 382 ACLPW--NFYNSSSTFEL-----VMSDFGVMTPIITVMI---CDNFLIRKQYSVTN 429
QY 323 GKCLWGRADSKIVASYTTTSGETKSLTSGM-----WGLARHFIYPEIL 370
Db 430 AFVL-----KGEYFTKGVNMRIV--AWVCGMADPLGIAMVNNDYFHTNGII 477
QY 371 SAFTWYPALEDFNLAIFYVIELTLLFD---RAKRDDECRKRGKY 415
Db 478 NFFYGD---SFSPLISFFYWGICLLPFRKIVKHDD---KQYGA 519

```

Search completed: January 14, 2003, 14:17:52

Job time : 26 secs

1	2280	87.9	430	1	US-08-601-435-2	Sequence 2, Appl
2	2280	87.9	430	2	US-08-601-447-2	Sequence 2, Appl
3	2280	87.9	430	2	US-08-783-202-2	Sequence 2, Appl
4	2280	87.9	430	4	US-09-443-041A-31	Sequence 31, Appl
5	1981	76.4	430	4	US-09-443-041A-18	Sequence 18, Appl
6	1976.5	76.2	450	4	US-09-443-041A-22	Sequence 22, Appl
7	1960	75.6	432	4	US-09-443-041A-20	Sequence 20, Appl
8	559	21.5	419	1	US-08-439-131A-3	Sequence 3, Appl
9	559	21.5	419	1	US-08-440-674-2	Sequence 3, Appl
10	554	21.4	134	4	US-09-443-041A-8	Sequence 8, Appl
11	544.5	21.0	161	4	US-09-443-041A-6	Sequence 6, Appl
12	499.5	19.3	418	4	US-09-442-653-7	Sequence 7, Appl
13	418.5	16.1	438	1	US-08-439-131A-2	Sequence 2, Appl
14	418.5	16.1	438	1	US-08-440-674-5	Sequence 5, Appl
15	368	14.2	473	1	US-08-439-131A-4	Sequence 4, Appl
16	368	14.2	473	1	US-08-440-674-3	Sequence 4, Appl
17	354.5	13.7	453	1	US-08-439-131A-5	Sequence 5, Appl
18	354.5	13.7	453	1	US-08-440-674-4	Sequence 4, Appl
19	253	9.8	74	4	US-09-443-041A-2	Sequence 2, Appl
20	249	9.6	67	4	US-09-443-041A-4	Sequence 4, Appl
21	216.5	8.3	145	4	US-09-442-653-4	Sequence 4, Appl
22	161	6.2	81	4	US-09-342-653-6	Sequence 6, Appl
23	105	4.0	68	4	US-09-342-653-2	Sequence 2, Appl
24	97	3.7	352	4	US-09-576-160B-3	Sequence 3, Appl
25	92.5	3.6	352	4	US-09-576-160B-1	Sequence 1, Appl
26	92	3.5	451	4	US-09-134-001C-4420	Sequence 4420, Appl
27	91.5	3.5	444	4	US-08-482-746-15	Sequence 15, Appl

45	81.5	3.1	252	4	US-09-134-001C-4109	Sequence 4109, App
44	82.5	3.2	1095	4	US-09-134-001C-4109	Sequence 4109, App
43	82.5	3.2	1095	4	US-09-134-001C-4109	Sequence 4109, App
42	82.5	3.2	467	4	US-09-134-001C-4109	Sequence 4109, App
41	83	3.2	3010	4	US-09-134-001C-4109	Sequence 4109, App
39	84	3.2	530	4	US-09-134-001C-4109	Sequence 4109, App
38	84.5	3.3	800	4	US-09-134-001C-4109	Sequence 4109, App
37	84.5	3.3	309	5	PCF-US93-08528-56	Sequence 56, App
36	85	3.3	446	4	US-08-833-696-11	Sequence 11, App
35	85	3.4	446	3	US-08-833-696-11	Sequence 11, App
34	87	3.4	446	3	US-08-833-696-11	Sequence 11, App
33	88	3.4	446	2	US-08-833-696-11	Sequence 11, App
32	88	3.4	446	2	US-08-833-696-11	Sequence 11, App
31	88	3.4	448	1	US-08-789-936-5	Sequence 5, App
30	88	3.4	448	1	US-08-789-936-5	Sequence 5, App
29	89.5	3.5	352	4	US-08-595-553A-2	Sequence 2, App
28	91.5	3.5	618	4	US-08-595-553A-2	Sequence 2, App

ALIGNMENTS

Db	1	MAETVSPITVATASMSLLIACFPFVILLWTYWHODGSVTOFFGFMNGVOGLINTWP	60
QY	1	MAETVSPITVATASMSLLIACFPFVILLWTYWHODGSVTOFFGFMNGVOGLINTWP	60
Db	1	MAETVSPITVATASMSLLIACFPFVILLWTYWHODGSVTOFFGFMNGVOGLINTWP	60
QY	61	RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVESPISPAGNRPVYKANGIAIYFTLATYIGL	120
Db	61	RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVESPISPAGNRPVYKANGIAIYFTLATYIGL	120
QY	121	WMGGINPAIVVDHGEIIFSAIIGSFIFCYLYIKGHVAPSSSDGSGCNLIIDFYWGM	180
Db	121	WMGGINPAIVVDHGEIIFSAIIGSFIFCYLYIKGHVAPSSSDGSGCNLIIDFYWGM	180


```

Db 181 ELVPRICKSDIKVTNCRGMSMAVLAVTYCIKOYEINGKVSMSLVNTIIMLVYTK 240
QY 241 FFWMEAGYWNMTMDIAHNRGCFYICMGCLVWVPSYTSFGMYLVNHPVELGTOLAIYTLVA 300
Db 241 FFWMEAGYWNMTMDIAHNRGCFYICMGCLVWVPSYTSFGMYLVNHPVELGTOLAIYTLVA 300
QY 301 GILCIYINYDCDRQROEFRTNKGCLWGRAPSKIVASYTTSGEYKTSLLTSGMWGLA 360
Db 301 GILCIYIKYDCDRQROEFRTNKGCLWGRAPSKIVASYTTSGEYKTSLLTSGMWGLA 360
QY 361 RHHYVEILISAFFWYTPALFDFNLAFYIYIETLLLFDRAKRDDRCRSKYKMYKLYC 420
Db 361 RHHYVEILISAFFWYTPALFDFNLAFYIYIETLLLFDRAKRDDRCRSKYKMYKLYC 418
QY 421 EK 422
Db 419 EK 420

RESULT 4
US-09-443-041A-31
; Sequence 31, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 31
; LENGTH: 430
; TYPE: PRF
; ORGANISM: Arabidopsis thaliana
US-09-443-041A-31

Query Match
Best Local Similarity 98.8%; Pred. No. 2, 7e-226;
Matches 417; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

```

```

QY 421 EK 422
Db 419 EK 420

RESULT 5
US-09-443-041A-18
; Sequence 18, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 18
; LENGTH: 430
; TYPE: PRF
; ORGANISM: Oryza sativa
US-09-443-041A-18

Query Match
Best Local Similarity 76.4%; Score 1981; DB 4; Length 430;
Best Local Similarity 83.3%; Pred. No. 1, 6e-195;
Matches 350; Conservative 34; Mismatches 36; Indels 0; Gaps 0;

```

```

:
: CURRENT APPLICATION NUMBER: US/09/443,041A
:
: CURRENT FILING DATE: 1999-11-18
:
: PRIOR APPLICATION NUMBER: 60/109,283
:
: PRIOR FILING DATE: 1998-11-20
:
: NUMBER OF SEQ ID NOS: 33
:
: SOFTWARE: Microsoft Office 97
:
: SEQ ID NO 22
:
: LENGTH: 450
:
: TYPE: PRT
:
: ORGANISM: Trifolium aestivum
:
: US-09-443-041A-22

```

Query Match	76.28;	Score 1976.5;	DB 4;	Length 450;
Best Local Similarity	83.68;	Pred. No. 4.9e-195;		
Matches 351;	Conservative 33;	Mismatches 35;	Indels 1;	Gaps 1;

[illegible]

```

US-09-443-041A-20
RESULT 7
: Sequence 20, Application US/09443041A
: Patent No. 6465717
: GENERAL INFORMATION:
: APPLICANT: Famodo, Omolayo O.
: APPLICANT: Oromoco, Buddy
: APPLICANT: Rafalski, Antoni
: APPLICANT: Shen, Jennie
: TITLE OF INVENTION: Sterol Metabolism Enzymes
: FILE REFERENCE:
: CURRENT APPLICATION NUMBER: US/09/443, 041A
: CURRENT FILING DATE: 1999-11-18
: PRIOR APPLICATION NUMBER: 60/109,283
: PRIOR FILING DATE: 1998-11-20
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: Microsoft Office 97
: SEQ ID NO 20
: LENGTH: 432
: TYPE: PRT
: ORGANISM: Glycine max
US-09-443-041A-20

```

Query Match	75.6%;	Score 1960;	DB 4;	Length 433;
Best Local Similarity	83.6%;	Pred. No. 2,3e-133;		
Matches 353;	Conservative 28;	Mismatches 41;	Indels 0;	Gaps 0

```

Db      1 MGNVTASPLVYASVISLTLCPFVLLMTWTMTADGSVSETHYKQNGLOGLHTWP 60
Qy      61 RPLVLAMKIIIFCYGAFEAIIOLLPRKRVBGRPSAGNRPVTKANGLAAYFVLTATYGL 120
Db      61 TPRPTACKIIAYAAFEAALOLLPRGTYYVGISPEGHHRYKANGLOAYFVLTITFEAL 120
Qy      121 WMGGINPAIYVDHGLGELFSALLIGSFPECVLLYXIGHNAPSSDSSGSGNLIIDFWGM 180
Db      121 WMGGINPPIYVHNHLEIYSALIGSFPECFVLYIGHNAPSSDSSGSGNLIIDFWGM 180
Qy      181 ELPRIGKSFEDIKVFETNCRFGMSMAVLAATYCCIKQYEINGKSDSMLVNTIMLVYTK 240
Db      181 ELPRIGKHFPIKVFETNCRFGMSMAVLAATYCCIKQYEINGKADBSMLVNTIMLVYTK 240
Qy      241 FFWMEAGYWNITDIAHDRGFGYICMGCLVWPVSVYTSPGMYLVNHPVELGTOLATYILVA 300
Db      241 FFWMEAGYWNITDIAHDRGFGYICMGCLVWPVSVYTSPGMYLVNHPVNLGKIALSTILVA 300
Qy      301 GIIICIIYNTDCRORORERRRRTNGKCLWGNAPRSKIYASTYTTSSGFKTSLTSSWMGGLA 360
Db      301 GIIICIIYNTDCRORORERRRRTNGKCLWGNAPRSKIEATYTTSSGFKTSLTSSWMGGLS 360
Qy      361 RHFHYPPELISAFFMTVVALADNPLAFYVYELTLLLFBRARDDDRCSKYGKWKLYC 420
Db      361 RHFHYPPELISAFMTVVALADNPLAFYVYELTLLLFBRARDDDRCSKYGKWKLYC 420
Qy      421 EK 422
Db      421 DK 422

```

RESULT 8
US-08-439-131A-3
: Sequence 3, Application US/08439131A
: Patent No. 5512472
: GENERAL INFORMATION:
: APPLICANT: Lai, Margaret H. K.
: APPLICANT: Bard, Martin
: APPLICANT: Kirsch, Donald R.
: TITLE OF INVENTION: A DNA Sequence Encoding Sterol Delta
: Patent No. 5512472
: TITLE OF INVENTION: Reductase
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: American Cyanamid Company
: STREET: One Cyanamid Plaza
: CITY: Wayne
: STATE: New Jersey
: COUNTRY: U.S.A.
: ZIP: 07470
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/439,131A
: FILING DATE: 11-MAY-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/107,347
: FILING DATE: 16-AUG-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Gordon, Alan M.
: REGISTRATION NUMBER: 30,637
: REFERENCE/DOCKET NUMBER: 854-012 (32,141)
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 201-831-3244
: TELEFAX: 201-831-3305
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 419 amino acids

```

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PUBLICATION INFORMATION:
AUTHORS: Worman, H. J.
AUTHORS: Evans, C. D.
AUTHORS: Blobel, G.
TITLE: The Lamin B Receptor of the Nuclear Envelope
JOURNAL: J. Cell Biol.
VOLUME: 111
PAGES: 1535-1542
DATE: 1990
RELEVANT RESIDUES IN SEQ ID NO: 3: FROM 190 TO 608
US-08-439-131A-3

Query Match
Best Local Similarity 21.5%; Score 559; DB 1; Length 419;
Matches 137; Conservative 62; Mismatches 178; Indels 40; Gaps 11;

QY 18 LLAFCPPFVILMYTHVHODGSVQTGF-----GFENGVGGLINIMPRPTLLAKKIIF 71
DB 21 MIMFPLPAIVLYLVLMCKODPSLMNFPPLPALESIMETKVFGLW-----FF 70
QY 72 CYGAFAEAILQLLPGRKVEG-PISPAGNRPVYKANGLAAYVTATYGLMWFGLFNPAT 130
DB 71 ----FOALFYLLPFGKVGVEGLPLSNP-RKPQYRINGFYAFLTAATQTLTYQ-FELHY 124
QY 131 VYDHGGEISALIFGSEIFCVLLYIKGHVAPSS--SDSGSCGNLIIDYWGMELYPRIG 188
DB 125 LHYHFOFAVSAAFSAFSAFSLIYIRSLKAPBEDLAPGNSGYLVYNEFGHELNPRIG- 183
QY 189 SFDIKVTNCRFGMSMAVLAATYTCIKOYEINGKYSDSM---LVNTILMYVYTKFEWME 245
DB 184 SFDIKVTNCRFGMSMAVLAATYTCIKOYEINGKYSDSM---LVNTILMYVYTKFEWME 242
QY 246 AGYWNMTDIAHNRGGEFYICWGLVWPVSYYTSPGMVLVNHPELGTQALAIYLVAGILCI 305
DB 243 EAVLTMTDITHDGFGLAFGLDLYWVPVYSLQAFYIVGHPIAISMPVAAITILNCIGY 302
QY 306 YINYDCDQROEFRRTNGKCLVWGAPSKIVASYTTTSGETKSLTLTSGWGLARHPHY 365
DB 303 YIFRSANSQKNFRN-----PADPKLSYLKVIPTATGKGLVTVGMGFEVHRHPY 352
QY 366 VPEILSAFMTVPALFDFNLAFYVIFLTLDFRAKRODDRCRSKYGKYLCKEK 422
DB 353 LGDITMALANSLPCGFNHLIPFYIYIFICLLVHREARDEHCKKRYGLAMERYCOR 409

RESULT 9
US-08-440-674-2
Sequence 2, Application US/08440674
Patent No. 5525496
GENERAL INFORMATION:
APPLICANT: Margaret H. Lai
TITLE OF INVENTION: A DNA Sequence Encoding Sterol
Patent No. 5525496
TITLE OF INVENTION: 14
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESS: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" 1.44 MB diskette
COMPUTER: IBM PC
OPERATING SYSTEM: MS DOS
SOFTWARE: Word Processor

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,674
FILING DATE: May 15, 1995
CLASSIFICATION: 536
PUBLICATION INFORMATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/107,347
FILING DATE: August 16, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Alan M. Gordon
REGISTRATION NUMBER: 30637
REFERENCE/DOCKET NUMBER: 854-P0012 DIV (32,141)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 419 residues
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
DESCRIPTION: polypeptide
FEATURE:
NAME/KEY: chicken
NAME/KEY: nuclear lamin B receptor
PUBLICATION INFORMATION:
AUTHORS: H. J. Worman, C. D. Evans, and G.
AUTHORS: Blobel
TITLE: (excerpt): The Lamin B Receptor of the
JOURNAL: Nuclear Envelope Inner Membrane
VOLUME: 111
PAGES: 1535-1542
PAGES: Sequence set out in Figure 5, page 1539
PATENT NO. 5525496
DATE: 1990
RELEVANT RESIDUES IN SEQ ID NO: 190 TO 608
US-08-440-674-2

Query Match
Best Local Similarity 21.5%; Score 559; DB 1; Length 419;
Matches 137; Conservative 62; Mismatches 178; Indels 40; Gaps 11;

QY 18 LLAFCPPFVILMYTHVHODGSVQTGF-----GFENGVGGLINIMPRPTLLAKKIIF 71
DB 21 MIMFPLPAIVLYLVLMCKODPSLMNFPPLPALESIMETKVFGLW-----FF 70
QY 72 CYGAFAEAILQLLPGRKVEG-PISPAGNRPVYKANGLAAYVTATYGLMWFGLFNPAT 130
DB 71 ----FOALFYLLPFGKVGVEGLPLSNP-RKPQYRINGFYAFLTAATQTLTYQ-FELHY 124
QY 131 VYDHGGEISALIFGSEIFCVLLYIKGHVAPSS--SDSGSCGNLIIDYWGMELYPRIG 188
DB 125 LHYHFOFAVSAAFSAFSAFSLIYIRSLKAPBEDLAPGNSGYLVYNEFGHELNPRIG- 183
QY 189 SFDIKVTNCRFGMSMAVLAATYTCIKOYEINGKYSDSM---LVNTILMYVYTKFEWME 245
DB 184 SFDIKVTNCRFGMSMAVLAATYTCIKOYEINGKYSDSM---LVNTILMYVYTKFEWME 242
QY 246 AGYWNMTDIAHNRGGEFYICWGLVWPVSYYTSPGMVLVNHPELGTQALAIYLVAGILCI 305
DB 243 EAVLTMTDITHDGFGLAFGLDLYWVPVYSLQAFYIVGHPIAISMPVAAITILNCIGY 302
QY 306 YINYDCDQROEFRRTNGKCLVWGAPSKIVASYTTTSGETKSLTLTSGWGLARHPHY 365
DB 303 YIFRSANSQKNFRN-----PADPKLSYLKVIPTATGKGLVTVGMGFEVHRHPY 352
QY 366 VPEILSAFMTVPALFDFNLAFYVIFLTLDFRAKRODDRCRSKYGKYLCKEK 422
DB 353 LGDITMALANSLPCGFNHLIPFYIYIFICLLVHREARDEHCKKRYGLAMERYCOR 409

RESULT 10

```

```

US-09-443-041A-8
; Sequence 8, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 134
; TYPE: PRF
; ORGANISM: Trifolium aestivum
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (124)
; OTHER INFORMATION: ANY AMINO ACID
US-09-443-041A-8

Query Match
21.4%; Score 554; DB 4; Length 134;
Best Local Similarity 75.4%; Pred. No. 2.3e-49;
Matches 101; Conservative 12; Mismatches 21; Indels 0; Gaps 0;

QY 4 TVHSPITVYASMSLSLAFCEPPEVILLMTVMHODGSVTQTFGEFMEVGOGLINIMPRT 63
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 TVHSAIVTYSMLALSLCPPEVILLMTVMHODGSVTRTEHLRDHGLEKAIWPMPT 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 64 LIAWKIIIFCGAFEAIIQLLPGRKEGPISPAGNRPVYKANGLAAYFTLATYGLMWF 123
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 LVAKKIIFGGLFEAVLQLLLPGRKEGPISPAGNRPVYKANGLAAYLATYGLMWF 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 124 GTFNPATVVDHGE 137
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 GTFNPATVVDHGE 134
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 11
US-09-443-041A-6
; Sequence 6, Application US/09443041A
; Patent No. 6465717
; GENERAL INFORMATION:
; APPLICANT: Farnodu, Omolayo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni
; APPLICANT: Shen, Jennie
; TITLE OF INVENTION: Sterol Metabolism Enzymes
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/443,041A
; CURRENT FILING DATE: 1999-11-18
; PRIOR APPLICATION NUMBER: 60/109,283
; PRIOR FILING DATE: 1998-11-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 6
; LENGTH: 161
; TYPE: PRF
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (77)
; OTHER INFORMATION: ANY AMINO ACID
; NAME/KEY: UNSURE
; LOCATION: (95)
; OTHER INFORMATION: ANY AMINO ACID
; NAME/KEY: UNSURE
; LOCATION: (102)

```

```

; OTHER INFORMATION: ANY AMINO ACID
; NAME/KEY: UNSURE
; LOCATION: (157)
; OTHER INFORMATION: ANY AMINO ACID
; NAME/KEY: UNSURE
; LOCATION: (159)
; OTHER INFORMATION: ANY AMINO ACID
US-09-443-041A-6

Query Match
21.0%; Score 544.5; DB 4; Length 161;
Best Local Similarity 64.8%; Pred. No. 2.9e-48;
Matches 105; Conservative 15; Mismatches 41; Indels 1; Gaps 1;

QY 1 MAETVHSPIVYASMSLSLAFCEPPEVILLMTVMHODGSVTQTFGEFMEVGOGLINIMP 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MGATVHSPIVYASVISLITLCPEFVILLMTVTLADGSVSEFTHYLRONGLOGLIWPT 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 RPTLIAMKIIIFCGAFEAIIQLLPGRKEGPISPAGNRPVYKANGLAAYFTLATYGL 120
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 TPTPTACKIIIVYAAFYAALQLLLPGRKEGPISPAGNRPVYKANGLAAYFTLATYFA- 119
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 WMFGTFNPATVVDHGEFSAIIIFGSPITFCVLYIKGHVAPS 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 WMFGTFNPATVVDHGEFSAIIIFGSPITFCVLYIKGHVAPS 161
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 12
US-09-342-653-7
; Sequence 7, Application US/09342653
; Patent No. 6306632
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Rafalski, Antoni
; TITLE OF INVENTION: Chromatin Associated Proteins
; FILE REFERENCE: BB-1118
; CURRENT APPLICATION NUMBER: US/09/342,653
; CURRENT FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: 60/092,841
; EARLIER FILING DATE: July 14, 1998
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 418
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-342-653-7

Query Match
19.3%; Score 499.5; DB 4; Length 418;
Best Local Similarity 32.8%; Pred. No. 4.8e-43;
Matches 123; Conservative 60; Mismatches 167; Indels 25; Gaps 8;

QY 54 GLINWPRPTLIAMKIIIFCGAFEAIIQLLPGRKEGPISPAGNRPVYKANGLAAYFT 113
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 53 GLEVIAMSPRALLW---LAWIGDQNALYLPARKVABGQELDKSRILRYPLNGQALVLT 109
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 114 LATYGLMWFGIFFNPATVVDHGEIFFSALIF---GSFIFCVLLIRGHVAPS--SDSG 167
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 110 -ALLVGLG---MSAGPLGALPEMLPLAVATLTAFIFELFLYMAQVAPASALAPGG 164
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 168 SCGNLIIDFWGMELYPRIKSPDIKVFYTNCRFGSMASVAVLATYCIKQYEINGKSDSM 227
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 165 NSGNPIYDFLGRLEINPRI-CEPFDEKYFCELRPGLIGVILNIALMLKKEELRSGPSLAM 223
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 228 LVNTILMLVYVYTKFFWAGVGNWMTMDLAHDGSGYIOMGCLVWAPSVYTSQWLVNHPV 287
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 224 WLWNGEQLLYVDALWBEAVLTTMDITHDGFMLAFGDAMVPFTYSIQAOPLHLHPQ 283
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 288 ELGTQLATYIIVAGILCIYINWCDROEFRRTNGKCLVWGRAPSKIVASYTTSETK 347
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 284 PLGLPMASVICLINALIGYIRFGANSQKNTRKKNPSDPRVAG-----LETISTATGRK- 336
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 348 TSLLTSGMWGLIARHFHYVPEILSAFEWTVPALFDNPLAVFYVLTLLLFDRRAKRODDR 407
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Query Match	16.1%;	Score 418.5;	DB 1;	Length 438,
Best Local Similarity	29.5%;	Pred. No. 1.1e-34;		

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 14:13:05 ; Search time 41 Seconds
(without alignments)
1547.007 Million cell updates/sec

Title: US-09-817-774-31
Perfect score: 2594
Sequence: 1 MAETVHSPIVTFYASMLSLA.....RAKTKMQDAIDILLICML 476

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: A.Geneseq-101002.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
24: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2312	89.1	432	21	Arabidopsis thaliana
2	2280	87.9	430	17	Arabidopsis thaliana
3	2242	86.4	418	21	Arabidopsis thaliana
4	2143	82.6	400	21	Arabidopsis thaliana
5	2011	77.5	376	21	Arabidopsis thaliana
6	1340	51.7	253	21	Arabidopsis thaliana
7	1220	47.0	231	21	Arabidopsis thaliana
8	698.5	26.9	475	20	Human secreted pro
9	698.5	26.9	475	20	Human Delta7-steroid
10	698.5	26.9	475	20	Human Delta7-steroid

11	698.5	26.9	475	22	AAU39059
12	698.5	26.9	475	22	ABR55768
13	651.5	25.1	154	21	AAAG40884
14	613.5	23.7	280	22	AAB60759
15	604.5	23.3	280	22	AAB60760
16	540	20.8	615	23	AAU84347
17	471.5	18.2	365	21	AAAG17152
18	471.5	18.2	365	21	AAAG17151
19	470.5	18.1	365	21	AAAG36707
20	470.5	18.1	369	19	AAW41576
21	470.5	18.1	369	21	AAAG38706
22	460	17.7	317	22	AAAB18063
23	419	16.2	374	22	AAAB20379
24	418.5	16.1	438	16	AAAT71934
25	418.5	16.1	438	17	AAAR96333
26	418.5	16.1	438	17	AAAW01975
27	405.5	15.6	369	22	AAAB20381
28	319.5	12.3	230	21	AAAG17153
29	318.5	12.3	230	21	AAAG38708
30	301	11.6	453	23	ABP41089
31	197	7.6	140	22	AAAB20380
32	173.5	6.7	716	22	ABBB6565
33	165	6.4	27	21	AAAB18126
34	104.5	4.0	423	22	AAAG93725
35	104.5	4.0	423	23	AAW49368
36	103	4.0	471	22	ABBS9886
37	103	4.0	500	22	ABBB67015
38	100	3.9	447	21	AAAG29468
39	100	3.9	447	21	ABBS34768
40	100	3.9	765	22	ABBB64843
41	100	3.9	765	22	AAAG30968
42	98	3.8	435	21	AAAG30967
43	98	3.8	479	21	AAAG30966
44	98	3.8	479	21	ABBB92289
45	98	3.8	1514	23	

ALIGNMENTS

RESULT 1	
AAAG53537	
ID	AAAG53537 standard; Protein; 432 AA.
XX	
AC	AAAG53537;
XX	
DT	18-OCT-2000 (first entry)
XX	
DE	Arabidopsis thaliana protein fragment SPQ ID NO: 68171.
XX	
KW	Protein identification; signal transduction pathway; metabolic pathway;
KW	hybridisation assay; genetic mapping; gene expression control; promoter;
KW	termination sequence.
XX	
OS	Arabidopsis thaliana.
XX	
PN	EP1033405-A2.
XX	
PD	06-SEP-2000.
XX	
PF	25-FEB-2000; 2000EP-0301439.
XX	
PR	25-FEB-1999; 99US-0121825.
PR	05-MAR-1999; 99US-0123180.
PR	09-MAR-1999; 99US-0123548.
PR	23-MAR-1999; 99US-0125788.
PR	25-MAR-1999; 99US-0126264.
PR	29-MAR-1999; 99US-0126785.
PR	01-APR-1999; 99US-0127462.
PR	06-APR-1999; 99US-0128234.
PR	08-APR-1999; 99US-0128714.
PR	16-APR-1999; 99US-0129845.
PR	19-APR-1999; 99US-0130077.

Human secreted pro
Human polypeptide
Zea mays protein f
Gene 13 related pe
Protein LBR differ
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis C-14 s
Pinus radiata ster
Corn sterol delta-ster
Sterol delta-14-re
Saccharomyces cere
Saccharomyces cere
Soybean sterol del
Arabidopsis thaliana
Arabidopsis thaliana
Human ovarian anti
Soybean sterol del
Drosophila melanog
Eucalyptus grandis
Human polypeptide
Human alcohol dehy
Drosophila melanog
Arabidopsis thaliana
Arabidopsis thaliana
Lactococcus lactis
Drosophila melanog
Arabidopsis thaliana
Arabidopsis thaliana
Arabidopsis thaliana
Herbicide activ

PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145224.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148685.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158223.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.

PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 89.1%; Score 2312; DB 21; Length 432;
 Best Local Similarity 99.8%; Pred. No. 4.6e-222;
 Matches 421; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAETVHSPIVTVASMLSLAFCEPPVILMTYMHODGSVYOTFGFWMNGVGLINIMP 60
 DB 1 MAETVHSPIVTVASMLSLAFCEPPVILMTYMHODGSVYOTFGFWMNGVGLINIMP 60
 QY 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVEGPISPAGNRPVYKANGLAAYVTLATYIGL 120
 DB 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVEGPISPAGNRPVYKANGLAAYVTLATYIGL 120
 QY 121 WMFGIENPAIVYDHLGEIFSAIIFGSEIFCYLLYIKGHVAPSSDSCGNLIDFYWGM 180
 DB 121 WMFGIENPAIVYDHLGEIFSAIIFGSEIFCYLLYIKGHVAPSSDSCGNLIDFYWGM 180
 QY 181 ELYPRIGKSPDIKFTNCRFGMSMAVLAAYCICKOYEINGKVSMSLVNTIMLVYVTK 240
 DB 181 ELYPRIGKSPDIKFTNCRFGMSMAVLAAYCICKOYEINGKVSMSLVNTIMLVYVTK 240
 QY 241 FFWMEAGYWNMTMDIAHNRGGEYICWGLVWPVSYSYTPGMVLVNHPELGTOLAIYILVA 300
 DB 241 FFWMEAGYWNMTMDIAHNRGGEYICWGLVWPVSYSYTPGMVLVNHPELGTOLAIYILVA 300
 QY 301 GILCIYINYDCDROREFRRTNGKCLVWGRAPSKIYASTTTSGEYKTSLLTSGMWGLA 360
 DB 301 GILCIYINYDCDROREFRRTNGKCLVWGRAPSKIYASTTTSGEYKTSLLTSGMWGLA 360
 QY 361 RHFHYVPEILSAFWTVPALFNDFLAYFYVIFLTLFDRAKRDDRCRSKYGKMYKLYC 420
 DB 361 RHFHYVPEILSAFWTVPALFNDFLAYFYVIFLTLFDRAKRDDRCRSKYGKMYKLYC 420
 QY 421 EK 422
 DB 421 EK 422

RESULT 2

AAW03567
 ID AAW03567 standard; Protein; 430 AA.

AC AAW03567;

DT 09-APR-1997 (first entry)

DE Arabidopsis thaliana delta-5,7-sterol, delta7-reductase.

XX Delta-7Red: 7-dehydrocholesterol reductase; C7-unsaturated sterol;
 KW pregnenolone; plant; delta-5,7 sterol, delta-7 reductase;

KW myastatin resistance.
 XX Arabidopsis thaliana.
 OS Arabidopsis thaliana.
 XX EP727489-A1.
 XX 21-AUG-1996.
 PD 14-FEB-1996; 96EP-0400301.
 PF 01-JUN-1995; 95FR-0006517.
 PR 15-FEB-1995; 95FR-0001723.
 XX (Rous) ROUSSEL-UCIAF.
 PA Chenivesse X, Dupont C, Lecain E, Pompon D;
 PI WPI; 1996-372876/38.
 DR N-PSDB; AAT39358.
 DR Nucleic acid encoding delta-5,7 sterol delta-7 reductase - esp. of
 PT A.thaliana, for producing pregnenolone-synthesizing yeast.
 PS Claim 6; Page 47-49; 82pp; French.

CC A myastatin-resistant clone containing cDNA which encodes the
 CC delta-5,7 sterol, delta-7 reductase enzyme of Arabidopsis thaliana
 CC was isolated from a cDNA library in yeast strain FY1679. The major
 CC sterol of the untransformed yeast strain is ergosterol. RP-HPLC
 CC analysis showed that in one clone the ergosterol was replaced by
 CC two major sterols which did not absorb at 285 nm. The insert from
 CC this clone was subcloned to localise the enzyme coding region and
 CC the cDNA sequence was determined. The enzyme having the present
 CC sequence is designated delta-7Red and reduces C7-unsaturated sterols.
 CC In particular, it is useful in the production of pregnenolone.

Sequence 430 AA:

Query Match 87.9%; Score 2280; DB 17; Length 430;
 Best Local Similarity 98.8%; Pred. No. 7.2e-215;
 Matches 417; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 MAETVHSPIVTVASMLSLAFCEPPVILMTYMHODGSVYOTFGFWMNGVGLINIMP 60
 DB 1 MAETVHSPIVTVASMLSLAFCEPPVILMTYMHODGSVYOTFGFWMNGVGLINIMP 60
 QY 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVEGPISPAGNRPVYKANGLAAYVTLATYIGL 120
 DB 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVEGPISPAGNRPVYKANGLAAYVTLATYIGL 120
 QY 121 WMFGIENPAIVYDHLGEIFSAIIFGSEIFCYLLYIKGHVAPSSDSCGNLIDFYWGM 180
 DB 121 WMFGIENPAIVYDHLGEIFSAIIFGSEIFCYLLYIKGHVAPSSDSCGNLIDFYWGM 180
 QY 181 ELYPRIGKSPDIKFTNCRFGMSMAVLAAYCICKOYEINGKVSMSLVNTIMLVYVTK 240
 DB 181 ELYPRIGKSPDIKFTNCRFGMSMAVLAAYCICKOYEINGKVSMSLVNTIMLVYVTK 240
 QY 241 FFWMEAGYWNMTMDIAHNRGGEYICWGLVWPVSYSYTPGMVLVNHPELGTOLAIYILVA 300
 DB 241 FFWMEAGYWNMTMDIAHNRGGEYICWGLVWPVSYSYTPGMVLVNHPELGTOLAIYILVA 300
 QY 301 GILCIYINYDCDROREFRRTNGKCLVWGRAPSKIYASTTTSGEYKTSLLTSGMWGLA 360
 DB 301 GILCIYINYDCDROREFRRTNGKCLVWGRAPSKIYASTTTSGEYKTSLLTSGMWGLA 360
 QY 361 RHFHYVPEILSAFWTVPALFNDFLAYFYVIFLTLFDRAKRDDRCRSKYGKMYKLYC 420
 DB 361 RHFHYVPEILSAFWTVPALFNDFLAYFYVIFLTLFDRAKRDDRCRSKYGKMYKLYC 420
 QY 421 EK 422
 DB 421 EK 422

RESULT 3
AAG53538
ID AAG53538 standard; Protein; 418 AA.
XX
AC AAG53538;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 68172.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127432.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145112.
PR 22-JUL-1999; 99US-0145115.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR	25-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139452.
PR	17-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	18-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142820.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	06-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159283.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	18-OCT-1999;	99US-0159568.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.

PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.
Query Match 82.6%; Score 2143; DB 21; Length 400;
Best Local Similarity 99.7%; Pred. No. 3,1e-205;
Matches 389; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 33 MVHDDGAVTQTFEFGFWENGVOGLINIMPRPLIMKTIIFCGAFETLLOLLPEKRVGP 92
Db 1 MVHDDGAVTQTFEFGFWENGVOGLINIMPRPLIMKTIIFCGAFETLLOLLPEKRVGP 60
QY 93 ISPGNRPVYKANGLAAYFVTLATYLGIMWGFIPALVYDHLGEISALIFGSFICVL 152
Db 61 ISPGNRPVYKANGLAAYFVTLATYLGIMWGFIPALVYDHLGEISALIFGSFICVL 120
QY 153 LYIKGHVAPSSSDSGCNLIIDFYKMELEYPRGKSFIDKVFNCGFGMSAVLAATY 212
Db 121 LYIKGHVAPSSSDSGCNLIIDFYKMELEYPRGKSFIDKVFNCGFGMSAVLAATY 180
QY 213 CIRQYEINGKVSMSLVNTILMYVTRKFWMEAGYWNMTDIADRGCFYICMGCLVWP 272
Db 181 CIRQYEINGKVSMSLVNTILMYVTRKFWMEAGYWNMTDIADRGCFYICMGCLVWP 240
QY 273 SVYSPGMYLVNHPVEIGTQAIYILVAGILCIYINDCDROQREFRRTNGKCLVWGRAP 332
Db 241 SVYSPGMYLVNHPVEIGTQAIYILVAGILCIYINDCDROQREFRRTNGKCLVWGRAP 300
QY 333 SKIYASTTSGEFTKSTLLTSSGMMGLARHHPYPELISAFPTVPLAFNGLAYFYVF 332
Db 301 SKIYASTTSGEFTKSTLLTSSGMMGLARHHPYPELISAFPTVPLAFNGLAYFYVF 360
QY 393 LTLLFDRAKRDDRCRSKYGKWKLYCEK 422
Db 361 LTLLFDRAKRDDRCRSKYGKWKLYCEK 390

RESULT 5
ID AAG23597 standard; Protein; 376 AA.
AC AAG23597;
DT 17-OCT-2000 (first entry)
DE Arabidopsis thaliana protein fragment SEQ ID NO: 26964.
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
PN EP1033405-A2.
PD 06-SEP-2000.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130049.
PR 23-APR-1999; 9905-0130510.
PR 28-APR-1999; 9905-0130891.
PR 28-APR-1999; 9905-0131449.

PR 30-APR-1999; 9905-0132048.
PR 30-APR-1999; 9905-0132407.
PR 04-MAY-1999; 9905-0132484.
PR 05-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 06-MAY-1999; 9905-0132487.
PR 07-MAY-1999; 9905-0132863.
PR 11-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.
PR 27-MAY-1999; 9905-0136392.
PR 28-MAY-1999; 9905-0136782.
PR 01-JUN-1999; 9905-0137222.
PR 03-JUN-1999; 9905-0137528.
PR 04-JUN-1999; 9905-0137502.
PR 07-JUN-1999; 9905-0137724.
PR 08-JUN-1999; 9905-0138094.
PR 10-JUN-1999; 9905-0138540.
PR 10-JUN-1999; 9905-0138847.
PR 14-JUN-1999; 9905-0139119.
PR 16-JUN-1999; 9905-0139452.
PR 16-JUN-1999; 9905-0139453.
PR 17-JUN-1999; 9905-0139492.
PR 18-JUN-1999; 9905-0139454.
PR 18-JUN-1999; 9905-0139455.
PR 18-JUN-1999; 9905-0139456.
PR 18-JUN-1999; 9905-0139457.
PR 18-JUN-1999; 9905-0139458.
PR 18-JUN-1999; 9905-0139459.
PR 18-JUN-1999; 9905-0139460.
PR 18-JUN-1999; 9905-0139461.
PR 18-JUN-1999; 9905-0139462.
PR 18-JUN-1999; 9905-0139463.
PR 18-JUN-1999; 9905-0139750.
PR 18-JUN-1999; 9905-0139763.
PR 21-JUN-1999; 9905-0139817.
PR 22-JUN-1999; 9905-0139899.
PR 23-JUN-1999; 9905-0140353.
PR 23-JUN-1999; 9905-0140354.
PR 24-JUN-1999; 9905-0140695.
PR 28-JUN-1999; 9905-0140823.
PR 29-JUN-1999; 9905-0140991.
PR 30-JUN-1999; 9905-0141287.
PR 01-JUL-1999; 9905-0141842.
PR 02-JUL-1999; 9905-0142154.
PR 06-JUL-1999; 9905-0142055.
PR 08-JUL-1999; 9905-0142390.
PR 09-JUL-1999; 9905-0142803.
PR 12-JUL-1999; 9905-0142920.
PR 13-JUL-1999; 9905-0142977.
PR 14-JUL-1999; 9905-0143542.
PR 15-JUL-1999; 9905-0143624.
PR 16-JUL-1999; 9905-0144005.
PR 16-JUL-1999; 9905-0144085.
PR 19-JUL-1999; 9905-0144086.
PR 19-JUL-1999; 9905-0144325.
PR 19-JUL-1999; 9905-0144331.
PR 19-JUL-1999; 9905-0144332.
PR 19-JUL-1999; 9905-0144333.
PR 19-JUL-1999; 9905-0144334.
PR 19-JUL-1999; 9905-0144335.
PR 20-JUL-1999; 9905-0144352.
PR 20-JUL-1999; 9905-0144632.
PR 20-JUL-1999; 9905-0144884.
PR 21-JUL-1999; 9905-0144814.

PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 04-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 05-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148365.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 23-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151338.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.

PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 26-OCT-1999; 99US-0161362.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 77.5% Score 2011; DB 21; Length 376;
 Best Local Similarity 99.7% Pred. No. 4,4e-192; Mismatches 1; Indels 0; Gaps 0;
 Matches 365; Conservative 0;

OY 57 NIMPRPTLIANKIIFCYGAFFAIILQLLPCKRVGSPISPAKNRPVYKANGIAAFVTLAT 116
 DB 1 NIMPRPTLIANKIIFCYGAFFAIILQLLPCKRVGSPISPAKNRPVYKANGIAAFVTLAT 60
 OY 117 YLGLMWEGIFNPAIVYDHLGEISFALIFGSPICVLLYIKGHVAPSSDSGSGNLIIDF 176
 DB 61 YLGLMWEGIFNPAIVYDHLGEISFALIFGSPICVLLYIKGHVAPSSDSGSGNLIIDF 120
 OY 177 YMGELVPRICKSPDIKFTNCRFGMSMAVLAATYCIKOYEINIGKXSDSMLVNTILMV 236
 DB 121 YMGELVPRICKSPDIKFTNCRFGMSMAVLAATYCIKOYEINIGKXSDSMLVNTILMV 180
 OY 237 YVTKFWMWAGYVNTMDIAHNRGSGFYICWGLVWPVSYSYSPKXVLNVHVELETOLAIY 296
 DB 181 YVTKFWMWAGYVNTMDIAHNRGSGFYICWGLVWPVSYSYSPKXVLNVHVELETOLAIY 240
 OY 297 ILVAGILCITYINPDORROEFRTNGKCLWGRAPSKIVASYTTTSGEIKTSILLTSGW 356
 DB 241 ILVAGILCITYINPDORROEFRTNGKCLWGRAPSKIVASYTTTSGEIKTSILLTSGW 300
 OY 357 WGLARHFHYVPEIISAFFWTPALFDNPLAFYVIFLTLFDRAKDDRCRSKYKYM 416
 DB 301 WGLARHFHYVPEIISAFFWTPALFDNPLAFYVIFLTLFDRAKDDRCRSKYKYM 360
 OY 417 KLYCEK 422
 DB 361 KLYCEK 366

RESULT 6
 AAG23598
 ID AAG23598 standard; Protein; 253 AA.
 XX AAG23598;
 AC
 XX
 DT 17-OCT-2000 (first entry)
 XX
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 26965.
 XX
 KM Protein identification: signal transduction pathway; metabolic pathway;
 KM hydriidation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 OS Arabidopsis thaliana.
 XX
 PM EPI033405-A2.
 XX
 PD 06-SEP-2000.
 XX

PR 25-FEB-2000; 2000EP-0301439.
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 23-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137528.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141887.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.

```

PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match
 Best Local Similarity 99.6%; Score 1340; DB 21; Length 253;
 Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 180 MELYPRIGKSFDIKYFTNCRFGMSWAVLAVTYCIKQYBINGKSVDSMLVNTLMVYVT 239
DB 1 MELYPRIGKSFDIKYFTNCRFGMSWAVLAVTYCIKQYBINGKSVDSMLVNTLMVYVT 60
QY 240 KFPWMEAGYWNMTMDIAHDGEPICWGLVWPSPYTSGLVNHVPVLGQLAIYIIV 299
DB 61 KFPWMEAGYWNMTMDIAHDGEPICWGLVWPSPYTSGLVNHVPVLGQLAIYIIV 120
QY 300 AGILCIYINWCDROHREFRRTNGKCLVWGRAPSKIVASYTTSGETKSLTSGMMGL 359
DB 121 AGILCIYINWCDROHREFRRTNGKCLVWGRAPSKIVASYTTSGETKSLTSGMMGL 180
QY 360 ARHFHVPELILSAFWTVALEDFNLAYFYVIFLTLLEFDRAKRDDRCRSRYKWKLY 419
DB 181 ARHFHVPELILSAFWTVALEDFNLAYFYVIFLTLLEFDRAKRDDRCRSRYKWKLY 240
QY 420 CEK 422
DB 241 CEK 243

```

RESULT 7
 ID AAG23599 standard; Protein; 231 AA.
 AC AAG23599;
 XX
 DT 17-OCT-2000 (first entry)
 XX Arabidopsis thaliana protein fragment SEQ ID NO: 26966.
 DE Protein identification; signal transduction pathway; metabolic pathway;
 XX
 KW

```

KW hybridisation assay; genetic mapping; gene expression control; promoter;  

KW termination sequence.  

XX Arabidopsis thaliana.  

OS EPI033405-A2.  

XX  

XX 06-SEP-2000.  

XX  

XX 25-FEB-2000; 2000EP-0301439.  

XX  

XX 25-FEB-1999; 99US-0121925.  

XX 05-MAR-1999; 99US-0123180.  

XX 09-MAR-1999; 99US-0123548.  

XX 23-MAR-1999; 99US-0125788.  

XX 25-MAR-1999; 99US-0126264.  

XX 29-MAR-1999; 99US-0126785.  

XX 01-APR-1999; 99US-0127462.  

XX 06-APR-1999; 99US-0128234.  

XX 08-APR-1999; 99US-0128714.  

XX 16-APR-1999; 99US-0129845.  

XX 19-APR-1999; 99US-0130077.  

XX 21-APR-1999; 99US-0130449.  

XX 23-APR-1999; 99US-0130510.  

XX 28-APR-1999; 99US-0130891.  

XX 30-APR-1999; 99US-0132048.  

XX 04-MAY-1999; 99US-0132407.  

XX 04-MAY-1999; 99US-0132484.  

XX 05-MAY-1999; 99US-0132485.  

XX 06-MAY-1999; 99US-0132486.  

XX 06-MAY-1999; 99US-0132487.  

XX 07-MAY-1999; 99US-0132863.  

XX 11-MAY-1999; 99US-0134256.  

XX 14-MAY-1999; 99US-0134218.  

XX 14-MAY-1999; 99US-0134219.  

XX 14-MAY-1999; 99US-0134221.  

XX 18-MAY-1999; 99US-0134370.  

XX 18-MAY-1999; 99US-0134768.  

XX 19-MAY-1999; 99US-0134941.  

XX 20-MAY-1999; 99US-0135124.  

XX 21-MAY-1999; 99US-0135353.  

XX 24-MAY-1999; 99US-0135629.  

XX 25-MAY-1999; 99US-0136021.  

XX 27-MAY-1999; 99US-0136392.  

XX 28-MAY-1999; 99US-0136782.  

XX 01-JUN-1999; 99US-0137222.  

XX 03-JUN-1999; 99US-0137528.  

XX 04-JUN-1999; 99US-0137502.  

XX 07-JUN-1999; 99US-0137724.  

XX 08-JUN-1999; 99US-0138094.  

XX 10-JUN-1999; 99US-0138540.  

XX 10-JUN-1999; 99US-0138847.  

XX 14-JUN-1999; 99US-0139119.  

XX 16-JUN-1999; 99US-0139452.  

XX 16-JUN-1999; 99US-0139453.  

XX 17-JUN-1999; 99US-0139492.  

XX 18-JUN-1999; 99US-0139454.  

XX 18-JUN-1999; 99US-0139455.  

XX 18-JUN-1999; 99US-0139456.  

XX 18-JUN-1999; 99US-0139457.  

XX 18-JUN-1999; 99US-0139457.  

XX 18-JUN-1999; 99US-0139458.  

XX 18-JUN-1999; 99US-0139459.  

XX 18-JUN-1999; 99US-0139460.  

XX 18-JUN-1999; 99US-0139461.  

XX 18-JUN-1999; 99US-0139462.  

XX 18-JUN-1999; 99US-0139463.  

XX 18-JUN-1999; 99US-0139750.  

XX 18-JUN-1999; 99US-0139763.  

XX 21-JUN-1999; 99US-0139817.  

XX 22-JUN-1999; 99US-0139899.  

XX 23-JUN-1999; 99US-0140353.  

XX 23-JUN-1999; 99US-0140354.

```


PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142350.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144682.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145319.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.

PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 13-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 47.0%; Score 1220; DB 21; Length 231;
Best Local Similarity 99.5%; Pred. No. 2,2e-113;
Matches 220; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 202 MMSMVLAVTYCIRKQYEINGKVS D S M L V N T I L M L V Y T K F F W M E G Y N T M D I A H D R G C F 261
DB 1 MMSMVLAVTYCIRKQYEINGKVS D S M L V N T I L M L V Y T K F F W M E G Y N T M D I A H D R G C F 60
QY 262 YICWGLVWVPSVYTS P G M Y L V N H P E L G T O L A I Y I L V A G I C I Y I N D C R O R O E F R T 321
DB 61 YICWGLVWVPSVYTS P G M Y L V N H P E L G T O L A I Y I L V A G I C I Y I N D C R O R O E F R T 120
QY 322 N K C L V W G R A P S K I V A S Y T T S G E F K T S L L T S G M W G L A R H F H Y P E I L S A F W T V P A L F 381
DB 121 N K C L V W G R A P S K I V A S Y T T S G E F K T S L L T S G M W G L A R H F H Y P E I L S A F W T V P A L F 180
QY 382 D N F L A V F Y I F L T L L P R A K R D D R C R S K Y G K Y K L C E K 422
DB 181 D N F L A V F Y I F L T L L P R A K R D D R C R S K Y G K Y K L C E K 221

RESULT 8
AAV29333
ID AAV29333 standard; Protein: 475 AA.
XX
AC AAV29333;
XX


```

QY 63 TLIAKTIICYGAFAILLQLLP-----GKRVGEGISPAGNRPVYKANGIAAYF 111
DQ 93 TRKAQOLTYLTWTFQVLYTSLPDCCHKFLPGYVGIOEGAVTPAGVNNKQIINGLOAML 152
QY 112 VTLATYTLGLMWF-----GIFNPAIVYDH-LGEIFSAALIFGSFICVLLYIKGHVAPSSS 164
DQ 153 LT-----HLMFANAHLLSWFSPTIIFDNMPLMCMANILG-YAVSTFAMWKGYFFPTSA 206
QY 165 -DSGSCGNLIDFYWMGMELYPRIGKSPDIKFTNCRFGMSMAVLAATYCIKOYEINGKY 223
DQ 207 RDCKFTEGNFNYMMGIEFNPRIGKWFDEKLFNGRPGIVAMTILINISPAKORELHSHV 266
QY 224 SDSMLVNTILMLVYTKFFWMEAGYWNMTDIAHDRGGEYICWGLVWVPSYTSFGVLY 283
DQ 267 TNAMVLVNLQAIYVIDEFWNETWYTKTIDICHDFGMYLGMGDCVWLPYLYTGLIY 326
QY 284 NHPVELGTOLAIYIIVAGIICITINYDCDROQROEFRTNKGCLVWGRAPSKIVASYTTTS 343
DQ 327 YHPVQLSTPHAVGVLLGLVGYIFRVAHQKDLFRRTDRCCLIMGRPKVIECSYTSAD 386
QY 344 GETKTSLLTSGMWGLARHFHYVEILSAFWTVPALDFNFAVYVIFLTLIPDRAKR 403
DQ 387 GQRHSHKLLVSGFWGVARHNENYVDLMGSLAYCLACGGHLLPFYITTYMALLTHRCR 446
QY 404 DDDRCRSKYGKWKLY 419
DQ 447 DEHRCAASKYGRDMERY 462

RESULT 10
ID AAW93573 standard; Protein; 475 AA.
AC AAW93573;
XX
DT 17-JUN-1999 (first entry)
XX
DE Human Delta7-sterol reductase protein.
XX
KW Human; Delta7-sterol reductase; hereditary; Smith-Lemli-Opitz syndrome;
KW diagnosis; screening; double bond removal; 7-dehydrocholesterol;
KW organic polymeric ring; cholesterol.
XX
OS Homo sapiens.
XX
PN DE19739940-A1.
XX
PD 18-MAR-1999.
XX
PE 11-SEP-1997; 97DE-1039940.
XX
PR 11-SEP-1997; 97DE-1039940.
XX
PA (GLOS/) GLOSSMANN H.
PI Fitzky B, Glossmann H, Moebius F;
XX
DR WPI; 1999-191430/17.
XX
N-PSDB; AAX23387.
XX
PT Human Delta7-sterol reductase polypeptide - useful for diagnosis or
PT treatment of genetic defects e.g. hereditary Smith-Lemli-Opitz
PT syndrome
XX
PS Disclosure; Page 29-30; 62pp; German.
XX
XX This invention describes a human Delta7-sterol reductase. The encoding
XX DNA can be used to diagnose or correct human Delta7-sterol reductase gene
XX defects e.g. hereditary Smith-Lemli-Opitz syndrome, and to produce the
XX recombinant Delta7-sterol reductase polypeptide, which can be used to
XX replace a defective Delta7-sterol reductase enzyme in humans or other
XX animals. It is also useful to screen for Delta7-sterol reductase
XX inhibitors or to introduce and remove double bonds in synthetic and

```

```

CC naturally occurring organic polymeric ring systems (Delta 7-sterol
CC reductase catalyses the conversion of 7-dehydrocholesterol to
CC cholesterol).
CC
XX
SQ Sequence 475 AA:
Query Match 26.9%; Score 698.5; DB 20; Length 475;
Best Local Similarity 35.3%; Pred. No. 7.2e-61;
Matches 154; Conservative 87; Mismatches 152; Indels 43; Gaps 10;

QY 13 ASMLSLAFCPPFVILLMTWVHOD-----GSYTGFGFWMGVGLINIM---PRP 62
DQ 41 ASYIFILLRNP---FVYVFIMACDYSCALGPVVDY-----TGHARLSDIMAKTPPI 92
QY 63 TLIAKTIICYGAFAILLQLLP-----GKRVGEGISPAGNRPVYKANGIAAYF 111
DQ 93 TRKAQOLTYLTWTFQVLYTSLPDCCHKFLPGYVGIOEGAVTPAGVNNKQIINGLOAML 152
QY 112 VTLATYTLGLMWF-----GIFNPAIVYDH-LGEIFSAALIFGSFICVLLYIKGHVAPSSS 164
DQ 153 LT-----HLMFANAHLLSWFSPTIIFDNMPLMCMANILG-YAVSTFAMWKGYFFPTSA 206
QY 165 -DSGSCGNLIDFYWMGMELYPRIGKSPDIKFTNCRFGMSMAVLAATYCIKOYEINGKY 223
DQ 207 RDCKFTEGNFNYMMGIEFNPRIGKWFDEKLFNGRPGIVAMTILINISPAKORELHSHV 266
QY 224 SDSMLVNTILMLVYTKFFWMEAGYWNMTDIAHDRGGEYICWGLVWVPSYTSFGVLY 283
DQ 267 TNAMVLVNLQAIYVIDEFWNETWYTKTIDICHDFGMYLGMGDCVWLPYLYTGLIY 326
QY 284 NHPVELGTOLAIYIIVAGIICITINYDCDROQROEFRTNKGCLVWGRAPSKIVASYTTTS 343
DQ 327 YHPVQLSTPHAVGVLLGLVGYIFRVAHQKDLFRRTDRCCLIMGRPKVIECSYTSAD 386
QY 344 GETKTSLLTSGMWGLARHFHYVEILSAFWTVPALDFNFAVYVIFLTLIPDRAKR 403
DQ 387 GQRHSHKLLVSGFWGVARHNENYVDLMGSLAYCLACGGHLLPFYITTYMALLTHRCR 446
QY 404 DDDRCRSKYGKWKLY 419
DQ 447 DEHRCAASKYGRDMERY 462

RESULT 11
ID AAW93059 standard; Protein; 475 AA.
AC AAW93059;
XX
DT 16-JAN-2002 (first entry)
XX
DE Human secreted protein pj323_2.
XX
XX Human; secreted protein; antiinflammatory; immunosuppressive;
XX neurotropic; neuroprotective; antiarthritic; antimicrobial; vulnerary;
XX cytosolic; antidiabetic; vitricide; antiferility; anticoagulant;
XX vasotropic; antiparkinsonian; immunostimulant; dermatological;
XX antitumoral; antitumor; antileuk; osteopathic; tranquiliser;
XX cerebroprotective; cytokine; cell proliferation; cell differentiation;
XX immune deficiency; severe combined immunodeficiency; SCID; tumour;
XX autoimmune disorder; multiple sclerosis; rheumatoid arthritis;
XX graft-versus-host disease; myeloid deficiency; wound healing; ulcer;
XX periodontal disease; osteoporosis; osteoarthritis; Alzheimer's disease;
XX Parkinson's disease; Huntington's disease; infection; cardiac disease;
XX stroke; sepsis; inflammatory bowel disease; contraceptive; immunogen;
XX food supplement; vaccine.
XX
XX Homo sapiens.
XX
XX WO200175068-A2.
XX
XX 11-OCT-2001.
XX

```


PT New secreted proteins and encoding polynucleotides, useful in gene
PT therapies, particularly for preventing or treating autoimmune
PT disorders, cancer, graft-versus-host disease, wound, osteoporosis,
PT stroke or inflammations

XX Disclosure; Page 281-282; 349pp; English.

XX The invention relates to isolated polynucleotides (ABA90876-ABA90968 and
CC ABA90980) and encoded proteins (ABAS55698-ABAS55800), especially
CC polynucleotides SEQ ID NO.1 (ABA90876) and SEQ ID NO.19 (ABA90885) and
CC proteins SEQ ID NO.2 (ABAS55698) and SEQ ID NO.20 (ABAS55707) contained in
CC clones b3306-7 and yb8-1 respectively and the clones b3306-7 and yb8-1
CC are deposited with the American Type Culture Collection (ATCC) with
CC accession number 98599. The polynucleotides and encoded polypeptides have
CC cytosolic, anti-inflammatory, immunomodulator, vulnerability,
CC neuroprotective, activin, inhibin, chemotactic, haemostatic, thrombolytic
CC and anti-inflammatory activity and acting as cytokine modulators,
CC hematopoiesis regulators, tissue growth modulators and/or cadherin
CC suppressors. The polypeptides and polynucleotides are useful in gene
CC therapies, particularly for preventing, treating or ameliorating any of
CC the following diseases: immune deficiency and disorders; e.g. bacterial
CC or fungal infections, autoimmune disorders, cancer, systemic lupus
CC erythematosus or graft-versus-host disease; myeloid or lymphoid cell
CC deficiencies; wound, burns, incisions and ulcers, osteoporosis or
CC osteoarthritis; central and peripheral nervous system diseases and
CC neuropathies, e.g. Alzheimer's, Parkinson's disease, Huntington's
CC disease, amyotrophic lateral sclerosis or Shy-Drager syndrome;
CC haemophilia, cardiac infarction or stroke; inflammations, shock, sepsis
CC or systemic inflammatory response syndrome, ischaemia-reperfusion
CC injury, endotoxin lethality, arthritis, inflammatory bowel disease or
CC Crohn's disease; or tumours or cancers, pemphigus vulgaris or pemphigus
CC foliaceus.

XX Sequence 475 AA:

Query Match 26.9%; Score 698.5; DB 23; Length 475;
Best Local Similarity 35.3%; Pred. No. 7.2e-61;
Matches 134; Conservative 87; Mismatches 152; Indels 43; Gaps 10;

QY 13 ASMSLAFPPFVILMTVMVHOD-----GSVOTCFEFTFENVOGLINTW---PPR 62
DB 41 ASYFILLFPP-----FIYIYIMCDQYSCALTPPVYDIY-----TGHARISDIWARTPTPI 92
QY 63 TLIAKKIIFCYGAFEALLQDLLP-----GKRVESGPISPAGNRPYKANGLAAYF 111
DB 93 TRKAAQLYTLAWTFQVLYLTLDPFCCKFLPGVVGIOBCAVTPAGVKNYQINGLOAMT 152
QY 112 VTILATYIGLWVF-----GIFNPAIYDH-LGEIFSAIFGSIFFCVLYIKGHVAPSS 164
DB 153 LT-----HILMFANAHLLSWFSPTIIFDMWIPILMCANILG-YAVSTFAMVKGFPPPTSA 206
QY 165 -DSGSGCNLIIDFYWGMEIYPRIGKSPDIKVFNTGCRFGMSMAVLAIVYCIKQYELNGVY 223
DB 207 RDKCFGNFNYNYMKGIEFNPRIKGFKMFDRKLFNGRPGVIAWTLINLSPAKGRELASHV 266
QY 224 SDSMLVNTILMIVYTKFEWMEAGVYNTMDIAHDSRGFYICWGLVWVSPVYSPGMIVY 283
DB 267 TNAMVAVNVLQAIYVIDFENETWYLTIDICHDRFGWLVGMDCWMLPYLYTLQGLYIV 326
QY 284 NHPEVETQAIATYIIVAGILCIYINDCDROKQEFRTNGKCLVWGRASKIYASTTTS 343
DB 327 YHPEVQSTPAVAVLLGLGVYIEFVANHOKDLFRTTRGKRLIWGRKPRVIECSTYSAD 386
QY 344 GETKTSLLTSGMWGLARHFHYVPELISAFFWTPALFNFNLAYFVIFLTLLEPRAR 403
DB 387 GQRHSHKLLVSGFWAGVARRHNHYGDLMSGLATYLAGGGHLLPYFIITMAILLTHRCIL 446
QY 404 DDDRCRSKYGKWKLY 419
DB 447 DEHRCAASKYGRDWERY 462

RESULT 13

AA040884
ID AA040884 standard; Protein; 154 AA.
XX
AC AA040884;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 50788.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
PN EP1033405-A2.
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0129845.
PR 21-APR-1999; 99US-0130077.
PR 23-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.

RESULT 14
AAB60759 standard; peptide: 280 AA.
ID AAB60759;
AC AAB60759;
XX
XX
XX 27-MAR-2001 (first entry)
XX
XX
XX Gene 13 related peptide #1.
XX
XX
XX Secreted protein; gene therapy; vaccine; cancer; leukemia;
XX autoimmune disease; allergy; inflammation; graft rejection;
XX hyperproliferation; cardiovascular; infection.
XX
XX Homo sapiens.
XX
XX WO200076531-A1.
XX
XX 21-DEC-2000.
XX
XX 01-JUN-2000; 2000WO-US15137.
XX
XX 11-JUN-1999; 99US-0138625.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis GA;
XX WPI; 2001-071148/08.
XX
XX Nucleic acids encoding 47 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX disease and diabetic retinopathy -
XX
XX Disclosure; Page 513-514; 525pp; English.
XX
XX The present invention relates to 26 secreted human proteins. The
XX proteins may be used in the prevention, diagnosis and treatment of
XX diseases associated with inappropriate polypeptide expression.
XX For example, they may be used in gene therapy or in vaccines.
XX Typical of diseases which are potentially treatable are cancers
XX (including leukemia), autoimmune diseases, allergies, inflammation,
XX graft rejection, hyperproliferation, cardiovascular diseases
XX (particularly critical limb ischemia and coronary disease) and any
XX involving abnormal angiogenesis, neurodegeneration and/or
XX infectious diseases.
XX
XX Sequence 280 AA:
SQ
Query Match 23.7%; Score 613.5; DB 22; Length 280;
Best Local Similarity 41.7%; Pred. No. 1.1e-52;
Matches 111; Conservative 61; Mismatches 93; Indels 1; Gaps 1;
QY 155 IKGIVAPSSS-DGSGCENLIDFYWGMELYPRIGKSFIDVFNCRFGMSNAVLAVTYCI 213
DB 2 VKGFEFTSARDCKFTGNFNYNMGIENRPRIGKMFDFLFGNGRGIYAATLNLISFA 61
QY 214 IKOYEINGKVSMSMLVNTILMLVYVTKFPMWEGYVNTMDIAHDSRGFYICWGLVWPSV 273
DB 62 AKQRELSHVTNANMLVNLVQAIYVIDFPMETWYLTIDICDHGKWLGMGDCWMLFYL 121
QY 274 VYSPGMYLVNHNVEELGTOLAIYILVAGLICITYINDCROEFRRTNGKCLWGRAPS 333
DB 122 LYTLOGLYLVYHVOULSTPHAVGVLLGLGVYIIFRVANHQKDLFRTDGRCLWGRKPK 181
QY 334 KIYASVTTSGETKSTLLTSGWMLARHRYVPELISAFWTVPALFDFNPLAFVFIPL 393
DB 182 VIECSYTSADGQRHSHKLLVSGFWGVARHFNVYGDLMGSLAYCLACGGHLLPYFIITM 241
QY 394 TLLFPAKRDDECRSKYGYKWKLY 419
DB 242 ALLTHRCLEDEHRCASKYGRDWERY 267

RESULT 15
AAB60760 standard; peptide: 280 AA.
ID AAB60760;
AC AAB60760;
XX
XX
XX 27-MAR-2001 (first entry)
XX
XX
XX Gene 13 related peptide #2.
XX
XX
XX Secreted protein; gene therapy; vaccine; cancer; leukemia;
XX autoimmune disease; allergy; inflammation; graft rejection;
XX hyperproliferation; cardiovascular; infection.
XX
XX Homo sapiens.
XX
XX WO200076531-A1.
XX
XX 21-DEC-2000.
XX
XX 01-JUN-2000; 2000WO-US15137.
XX
XX 11-JUN-1999; 99US-0138625.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis GA;
XX WPI; 2001-071148/08.
XX
XX Nucleic acids encoding 47 human secreted polypeptides, useful for
XX preventing, diagnosing and/or treating e.g. cancers, Parkinson's
XX disease and diabetic retinopathy -
XX
XX Disclosure; Page 514-515; 525pp; English.
XX
XX The present invention relates to 26 secreted human proteins. The
XX proteins may be used in the prevention, diagnosis and treatment of
XX diseases associated with inappropriate polypeptide expression.
XX For example, they may be used in gene therapy or in vaccines.
XX Typical of diseases which are potentially treatable are cancers
XX (including leukemia), autoimmune diseases, allergies, inflammation,
XX graft rejection, hyperproliferation, cardiovascular diseases
XX (particularly critical limb ischemia and coronary disease) and any
XX involving abnormal angiogenesis, neurodegeneration and/or
XX infectious diseases.
XX
XX Sequence 280 AA:
SQ
Query Match 23.3%; Score 604.5; DB 22; Length 280;
Best Local Similarity 41.5%; Pred. No. 8.4e-52;
Matches 110; Conservative 60; Mismatches 94; Indels 1; Gaps 1;
QY 156 KGVAPSSS-DGSGCENLIDFYWGMELYPRIGKSFIDVFNCRFGMSNAVLAVTYCI 214
DB 3 KGVFEFTSARDCKFTGNFNYNMGIENRPRIGKMFDFLFGNGRGIYAATLNLISFA 62
QY 215 IKOYEINGKVSMSMLVNTILMLVYVTKFPMWEGYVNTMDIAHDSRGFYICWGLVWPSV 274
DB 63 KXRELSHVTNANMLVNLVQAIYVIDFPMETWYLTIDICDHGKWLGMGDCWMLFYL 122
QY 275 VYSPGMYLVNHNVEELGTOLAIYILVAGLICITYINDCROEFRRTNGKCLWGRAPS 334
DB 123 LYTLOGLYLVYHVOULSTPHAVGVLLGLGVYIIFRVANHQKDLFRTDGRCLWGRKPK 182
QY 335 KIYASVTTSGETKSTLLTSGWMLARHRYVPELISAFWTVPALFDFNPLAFVFIPL 394
DB 183 IECSTYTSADGQRHSHKLLVSGFWGVARHFNVYGDLMGSLAYCLACGGHLLPYFIITM 242
QY 395 TLLFPAKRDDECRSKYGYKWKLY 419
DB 243 ALLTHRCLEDEHRCASKYGRDWERY 267

Fri Jan 17 19:40:53 2003

us-09-817-774-31.rag

Page 18

Search completed: January 14, 2003, 14:15:10
Job time : 43 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 14, 2003, 14:17:29 ; Search time 13 Seconds
(without alignments)
710.377 Million cell updates/sec

Title: US-09-817-774-31

Perfect score: 2594

Sequence: 1 MAETVHSPIVTYASMLSLA.....RAKTKMQDAIDILLICML 476

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2594	100.0	476	10	US-09-817-774-31
2	2318	89.4	432	10	US-09-817-774-29
3	2313	89.2	432	10	US-09-817-774-39
4	2312	89.1	432	10	US-09-817-774-43
5	2247	86.6	411	10	US-09-817-774-45
6	2233	86.1	408	10	US-09-817-774-41
7	2182	84.1	399	10	US-09-817-774-37
8	1606	61.9	328	10	US-09-817-774-33
9	1606	61.9	323	10	US-09-817-774-35
10	710.5	27.4	471	10	US-09-817-774-27
11	698.5	26.9	473	10	US-09-729-674-12
12	698.5	26.9	475	10	US-09-817-774-26
13	553.5	21.3	427	10	US-09-817-774-23
14	540	20.8	426	10	US-09-817-774-22
15	500.5	19.3	424	10	US-09-817-774-25
16	97	3.7	615	9	US-09-738-626-6967
17	95.5	3.7	541	9	US-10-028-072-14
18	95.5	3.7	555	10	US-09-925-297-827
19	95.5	3.7	733	12	US-10-036-328A-4

20	95.5	3.7	1210	12	US-10-036-328A-2	Sequence 2, Appl1
21	94.5	3.6	500	12	US-10-033-109-4	Sequence 4, Appl1
22	94.5	3.6	541	10	US-09-809-345A-33	Sequence 33, Appl1
23	93	3.6	427	10	US-09-922-501-2	Sequence 2, Appl1
24	91.5	3.5	444	10	US-09-853-386-132	Sequence 132, Appl1
25	91.5	3.5	444	10	US-09-853-386-136	Sequence 136, Appl1
26	91.5	3.5	444	10	US-09-191-724-15	Sequence 15, Appl1
27	91.5	3.5	618	10	US-09-995-007-2	Sequence 2, Appl1
28	89.5	3.5	440	9	US-09-738-626-4646	Sequence 4646, Ap
29	89	3.4	420	10	US-09-853-386-129	Sequence 129, Ap
30	89	3.4	427	10	US-09-922-501-6	Sequence 6, Appl1
31	88	3.4	448	12	US-10-029-756-5	Sequence 5, Appl1
32	87.5	3.4	494	12	US-10-033-109-6	Sequence 6, Appl1
33	87.5	3.4	662	12	US-10-036-328A-8	Sequence 8, Appl1
34	87.5	3.4	714	10	US-09-818-264-2	Sequence 2, Appl1
35	87.5	3.4	1138	12	US-10-036-328A-6	Sequence 6, Appl1
36	87	3.4	259	10	US-09-905-176-20	Sequence 20, Appl1
37	86	3.3	526	9	US-09-738-626-5712	Sequence 5712, Ap
38	85.5	3.3	501	8	US-08-635-967-2	Sequence 2, Appl1
39	85.5	3.3	539	10	US-09-815-242-5139	Sequence 5139, Ap
40	85	3.3	446	9	US-10-136-517-19	Sequence 19, Appl1
41	84	3.2	390	10	US-09-925-302-661	Sequence 661, Ap
42	83.5	3.2	519	10	US-09-825-012-80	Sequence 80, Appl1
43	83	3.2	245	9	US-09-895-913A-48	Sequence 48, Appl1
44	83	3.2	321	10	US-09-886-055-123	Sequence 123, App
45	83	3.2	415	10	US-09-853-386-128	Sequence 128, App

ALIGNMENTS

RESULT 1
US-09-817-774-31
; Sequence 31, Application US/09817774
; Patent No. US2002012011A1
; GENERAL INFORMATION:
; APPLICANT: CHOE, Sunghwa
; APPLICANT: FELDMANN A., Kenneth
; TITLE OF INVENTION: DWT5 MUTANTS
; FILE REFERENCE: 2225-0020 / 91020.002
; CURRENT APPLICATION NUMBER: US/09/817, 774
; CURRENT FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/192, 202
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 31
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-817-774-31

Query Match 100.0%; Score 2594; DB 10; Length 476;
Best Local Similarity 100.0%; Pred. No. 4.8e-236;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAETVHSPIVTYASMLSLAFCPPVILLMTVMHQDSVTQTEGFEWENGVOGLINIMP 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 1 MAETVHSPIVTYASMLSLAFCPPVILLMTVMHQDSVTQTEGFEWENGVOGLINIMP 60
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 61 RPTLAWKTIIFCYGAFEMILLOLLPGKRVGEGISAGNRPPYKANGLAAYVTATATYIGL 120
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 WMFGIFENPAIVVDHGEIIFSAIFGSEFCVLLYTKGVAVASSSDSGSGMLIIDFYGM 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 121 WMFGIFENPAIVVDHGEIIFSAIFGSEFCVLLYTKGVAVASSSDSGSGMLIIDFYGM 180
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 ELYPRIGSFEDIKVTTCNCFGMSAVLAVTYCIKQYEINGKVSDSMLVNTILMLVYVTK 240
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 181 ELYPRIGSFEDIKVTTCNCFGMSAVLAVTYCIKQYEINGKVSDSMLVNTILMLVYVTK 240
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
QY 241 FFWMEAGYWNIMDIADHDGSGFICWGLVWVPSVYTSFGMTLVNHPVBLGTQLAIYIILVA 300

```

Db 241 FFWMEAGYWNMTMDIAHNRGFIYICWGLVWVPSVYTSFGMYLVNHPVELGTOLAIYILVA 300
QY 301 GILCIYINYDCDRQROEFRTNKGCLVWGRAPSKIVASTTTSGEKTSLILTSGMWGLA 360
Db 301 GILCIYINYDCDRQROEFRTNKGCLVWGRAPSKIVASTTTSGEKTSLILTSGMWGLA 360
QY 361 RHHYVPELISAFETWVPALFDNFLAYFYVIFLTLLEPRAKRDDRCRSKYGKWKLYC 420
Db 361 RHHYVPELISAFETWVPALFDNFLAYFYVIFLTLLEPRAKRDDRCRSKYGKWKLYC 420
QY 421 EKSNTGSFREFIDCNEVCSSHLLITLITLESSEKDRAKTKQIDAIDILLICML 476
Db 421 EKSNTGSFREFIDCNEVCSSHLLITLITLESSEKDRAKTKQIDAIDILLICML 476

```

RESULT 2

```

US-09-817-774-29
; Sequence 29, Application US/09817774
; Patent No. US2002012011A1
; GENERAL INFORMATION:
; APPLICANT: CHOE, Sunghwa
; APPLICANT: FELDMANN A., Kenneth
; TITLE OF INVENTION: Dwt5 MUTANTS
; FILE REFERENCE: 2225-0020 / 91020.002
; CURRENT APPLICATION NUMBER: US/09/817,774
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/192,202
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 432
; TYPE: PRF
; ORGANISM: Arabidopsis thaliana
US-09-817-774-29

```

Query Match 89.4%; Score 2318; DB 10; Length 432;

Best Local Similarity 100.0%; Pred. No. 3.7e-210; Matches 422; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAETHSPITVYASMLSLAFCEPPEVILLMTYVNHODGSVTQTFGEFWENGVOGLINIMP 60
Db 1 MAETHSPITVYASMLSLAFCEPPEVILLMTYVNHODGSVTQTFGEFWENGVOGLINIMP 60
QY 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVESGPISPAENRPPYKANGLAAYFVLATYGL 120
Db 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVESGPISPAENRPPYKANGLAAYFVLATYGL 120
QY 121 WMEGIFNPAIYVDHIGEFISALIFGSEIFCVLLYIKGHVAPSSSDSGSCGNLIIDFYWGM 180
Db 121 WMEGIFNPAIYVDHIGEFISALIFGSEIFCVLLYIKGHVAPSSSDSGSCGNLIIDFYWGM 180
QY 181 ELIPRIGKSPDIKVTNCRFGMSMAVLAATYCIKOYEINKEVSDSMVNTIIMLYVYTK 240
Db 181 ELIPRIGKSPDIKVTNCRFGMSMAVLAATYCIKOYEINKEVSDSMVNTIIMLYVYTK 240
QY 241 FFWMEAGYWNMTMDIAHNRGFIYICWGLVWVPSVYTSFGMYLVNHPVELGTOLAIYILVA 300
Db 241 FFWMEAGYWNMTMDIAHNRGFIYICWGLVWVPSVYTSFGMYLVNHPVELGTOLAIYILVA 300
QY 301 GILCIYINYDCDRQROEFRTNKGCLVWGRAPSKIVASTTTSGEKTSLILTSGMWGLA 360
Db 301 GILCIYINYDCDRQROEFRTNKGCLVWGRAPSKIVASTTTSGEKTSLILTSGMWGLA 360
QY 361 RHHYVPELISAFETWVPALFDNFLAYFYVIFLTLLEPRAKRDDRCRSKYGKWKLYC 420
Db 361 RHHYVPELISAFETWVPALFDNFLAYFYVIFLTLLEPRAKRDDRCRSKYGKWKLYC 420
QY 421 EK 422
Db 421 EK 422

```

RESULT 3

```

US-09-817-774-39
; Sequence 39, Application US/09817774
; Patent No. US2002012011A1
; GENERAL INFORMATION:
; APPLICANT: CHOE, Sunghwa
; APPLICANT: FELDMANN A., Kenneth
; TITLE OF INVENTION: Dwt5 MUTANTS
; FILE REFERENCE: 2225-0020 / 91020.002
; CURRENT APPLICATION NUMBER: US/09/817,774
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/192,202
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 432
; TYPE: PRF
; ORGANISM: Arabidopsis thaliana
US-09-817-774-39

```

Query Match 89.2%; Score 2313; DB 10; Length 432;

Best Local Similarity 99.8%; Pred. No. 1.1e-209; Matches 421; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAETHSPITVYASMLSLAFCEPPEVILLMTYVNHODGSVTQTFGEFWENGVOGLINIMP 60
Db 1 MAETHSPITVYASMLSLAFCEPPEVILLMTYVNHODGSVTQTFGEFWENGVOGLINIMP 60
QY 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVESGPISPAENRPPYKANGLAAYFVLATYGL 120
Db 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVESGPISPAENRPPYKANGLAAYFVLATYGL 120
QY 121 WMEGIFNPAIYVDHIGEFISALIFGSEIFCVLLYIKGHVAPSSSDSGSCGNLIIDFYWGM 180
Db 121 WMEGIFNPAIYVDHIGEFISALIFGSEIFCVLLYIKGHVAPSSSDSGSCGNLIIDFYWGM 180
QY 181 ELIPRIGKSPDIKVTNCRFGMSMAVLAATYCIKOYEINKEVSDSMVNTIIMLYVYTK 240
Db 181 ELIPRIGKSPDIKVTNCRFGMSMAVLAATYCIKOYEINKEVSDSMVNTIIMLYVYTK 240
QY 241 FFWMEAGYWNMTMDIAHNRGFIYICWGLVWVPSVYTSFGMYLVNHPVELGTOLAIYILVA 300
Db 241 FFWMEAGYWNMTMDIAHNRGFIYICWGLVWVPSVYTSFGMYLVNHPVELGTOLAIYILVA 300
QY 301 GILCIYINYDCDRQROEFRTNKGCLVWGRAPSKIVASTTTSGEKTSLILTSGMWGLA 360
Db 301 GILCIYINYDCDRQROEFRTNKGCLVWGRAPSKIVASTTTSGEKTSLILTSGMWGLA 360
QY 361 RHHYVPELISAFETWVPALFDNFLAYFYVIFLTLLEPRAKRDDRCRSKYGKWKLYC 420
Db 361 RHHYVPELISAFETWVPALFDNFLAYFYVIFLTLLEPRAKRDDRCRSKYGKWKLYC 420
QY 421 EK 422
Db 421 EK 422

```

SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 45
 LENGTH: 432
 TYPE: PRN
 ORGANISM: Arabidopsis thaliana
 US-09-817-774-45

Query Match 89.1%; Score 2312; DB 10; Length 432;
 Best Local Similarity 99.8%; Pred. No. 1.3e-209;
 Matches 421; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAETVSPITVYASMSLAFCEPPVILLMTYVHODGSVOTGFFGFWENGVOGLINIMP 60
 DB 1 MAETVSPITVYASMSLAFCEPPVILLMTYVHODGSVOTGFFGFWENGVOGLINIMP 60
 QY 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVGSPISPAKNRPYKANGLAAYVTLATYGL 120
 DB 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVGSPISPAKNRPYKANGLAAYVTLATYGL 120
 QY 121 WMFGIFNPAIVYDHLGEIFSAIIFGSPFCVLLYIKGHVAPSSDSCGNLIIDFYWGM 180
 DB 121 WMFGIFNPAIVYDHLGEIFSAIIFGSPFCVLLYIKGHVAPSSDSCGNLIIDFYWGM 180
 QY 121 WMFGIFNPAIVYDHLGEIFSAIIFGSPFCVLLYIKGHVAPSSDSCGNLIIDFYWGM 180
 DB 121 WMFGIFNPAIVYDHLGEIFSAIIFGSPFCVLLYIKGHVAPSSDSCGNLIIDFYWGM 180
 QY 181 ELPRIGKSEFDIKVTNCRFGMSMAVLAAYTCIKOYEINGKVSMSLVNTILMYVTK 240
 DB 181 ELPRIGKSEFDIKVTNCRFGMSMAVLAAYTCIKOYEINGKVSMSLVNTILMYVTK 240
 QY 241 FFWMEAGYWNMTMDIAHDSRGFYICWGLVWPVSPTSFGMYLVNHPVELGTOLATYILA 300
 DB 241 FFWMEAGYWNMTMDIAHDSRGFYICWGLVWPVSPTSFGMYLVNHPVELGTOLATYILA 300
 QY 301 GILCIYINDCDROQROEFRTNGKCLVWGRAPSKIVASYTTSGEFTKSLTSGMWGLA 360
 DB 301 GILCIYINDCDROQROEFRTNGKCLVWGRAPSKIVASYTTSGEFTKSLTSGMWGLA 360
 QY 361 RHRYVPELISAFETVPALFDFNLAIFYVIFLTLFDRAKDDRCRSKYKWKLYC 420
 DB 361 RHRYVPELISAFETVPALFDFNLAIFYVIFLTLFDRAKDDRCRSKYKWKLYC 420
 QY 421 EK 422
 DB 421 EK 422

RESULT 5
 US-09-817-774-43
 Sequence 43, Application US/09817774
 Patent No. US2002012011A1
 GENERAL INFORMATION:
 APPLICANT: CHOE, Sunghwa
 APPLICANT: FELDMANN A., Kenneth
 TITLE OF INVENTION: Dwf5 MUTANTS
 FILE REFERENCE: 2225-0020 / 91020.002
 CURRENT APPLICATION NUMBER: US/09/817,774
 CURRENT FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: 60/192,202
 PRIOR FILING DATE: 2000-03-27
 NUMBER OF SEQ ID NOS: 45
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 43
 LENGTH: 411
 TYPE: PRN
 ORGANISM: Arabidopsis thaliana
 US-09-817-774-43

Query Match 86.6%; Score 2247; DB 10; Length 411;
 Best Local Similarity 100.0%; Pred. No. 1.6e-203;
 Matches 411; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAETVSPITVYASMSLAFCEPPVILLMTYVHODGSVOTGFFGFWENGVOGLINIMP 60
 DB 1 MAETVSPITVYASMSLAFCEPPVILLMTYVHODGSVOTGFFGFWENGVOGLINIMP 60

QY 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVGSPISPAKNRPYKANGLAAYVTLATYGL 120
 DB 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVGSPISPAKNRPYKANGLAAYVTLATYGL 120
 QY 121 WMFGIFNPAIVYDHLGEIFSAIIFGSPFCVLLYIKGHVAPSSDSCGNLIIDFYWGM 180
 DB 121 WMFGIFNPAIVYDHLGEIFSAIIFGSPFCVLLYIKGHVAPSSDSCGNLIIDFYWGM 180
 QY 181 ELPRIGKSEFDIKVTNCRFGMSMAVLAAYTCIKOYEINGKVSMSLVNTILMYVTK 240
 DB 181 ELPRIGKSEFDIKVTNCRFGMSMAVLAAYTCIKOYEINGKVSMSLVNTILMYVTK 240
 QY 241 FFWMEAGYWNMTMDIAHDSRGFYICWGLVWPVSPTSFGMYLVNHPVELGTOLATYILA 300
 DB 241 FFWMEAGYWNMTMDIAHDSRGFYICWGLVWPVSPTSFGMYLVNHPVELGTOLATYILA 300
 QY 301 GILCIYINDCDROQROEFRTNGKCLVWGRAPSKIVASYTTSGEFTKSLTSGMWGLA 360
 DB 301 GILCIYINDCDROQROEFRTNGKCLVWGRAPSKIVASYTTSGEFTKSLTSGMWGLA 360
 QY 361 RHRYVPELISAFETVPALFDFNLAIFYVIFLTLFDRAKDDRCRSK 411
 DB 361 RHRYVPELISAFETVPALFDFNLAIFYVIFLTLFDRAKDDRCRSK 411

RESULT 6
 US-09-817-774-41
 Sequence 41, Application US/09817774
 Patent No. US2002012011A1
 GENERAL INFORMATION:
 APPLICANT: CHOE, Sunghwa
 APPLICANT: FELDMANN A., Kenneth
 TITLE OF INVENTION: Dwf5 MUTANTS
 FILE REFERENCE: 2225-0020 / 91020.002
 CURRENT APPLICATION NUMBER: US/09/817,774
 CURRENT FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: 60/192,202
 PRIOR FILING DATE: 2000-03-27
 NUMBER OF SEQ ID NOS: 45
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 41
 LENGTH: 408
 TYPE: PRN
 ORGANISM: Arabidopsis thaliana
 US-09-817-774-41

Query Match 86.1%; Score 2233; DB 10; Length 408;
 Best Local Similarity 100.0%; Pred. No. 3.3e-202;
 Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAETVSPITVYASMSLAFCEPPVILLMTYVHODGSVOTGFFGFWENGVOGLINIMP 60
 DB 1 MAETVSPITVYASMSLAFCEPPVILLMTYVHODGSVOTGFFGFWENGVOGLINIMP 60
 QY 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVGSPISPAKNRPYKANGLAAYVTLATYGL 120
 DB 61 RPTLIAMKIIIFCYGAFEAIIQLLLPGKRVGSPISPAKNRPYKANGLAAYVTLATYGL 120
 QY 121 WMFGIFNPAIVYDHLGEIFSAIIFGSPFCVLLYIKGHVAPSSDSCGNLIIDFYWGM 180
 DB 121 WMFGIFNPAIVYDHLGEIFSAIIFGSPFCVLLYIKGHVAPSSDSCGNLIIDFYWGM 180
 QY 121 WMFGIFNPAIVYDHLGEIFSAIIFGSPFCVLLYIKGHVAPSSDSCGNLIIDFYWGM 180
 DB 121 WMFGIFNPAIVYDHLGEIFSAIIFGSPFCVLLYIKGHVAPSSDSCGNLIIDFYWGM 180
 QY 181 ELPRIGKSEFDIKVTNCRFGMSMAVLAAYTCIKOYEINGKVSMSLVNTILMYVTK 240
 DB 181 ELPRIGKSEFDIKVTNCRFGMSMAVLAAYTCIKOYEINGKVSMSLVNTILMYVTK 240
 QY 241 FFWMEAGYWNMTMDIAHDSRGFYICWGLVWPVSPTSFGMYLVNHPVELGTOLATYILA 300
 DB 241 FFWMEAGYWNMTMDIAHDSRGFYICWGLVWPVSPTSFGMYLVNHPVELGTOLATYILA 300
 QY 301 GILCIYINDCDROQROEFRTNGKCLVWGRAPSKIVASYTTSGEFTKSLTSGMWGLA 360
 DB 301 GILCIYINDCDROQROEFRTNGKCLVWGRAPSKIVASYTTSGEFTKSLTSGMWGLA 360

QY 361 RHFHVPEILSAFWTVPALFQNFLEYVIFLTLLEPDRAKDDDC 408
 DB 361 RHFHVPEILSAFWTVPALFQNFLEYVIFLTLLEPDRAKDDDC 408

RESULT 7

US-09-817-774-37
 ; Sequence 37, Application US/09817774
 ; Patent No. US2002012011A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHOE, Sunghwa
 ; APPLICANT: FELDMANN A., Kenneth
 ; TITLE OF INVENTION: Dwf5 MUTANTS
 ; FILE REFERENCE: 2225-0020 / 91020.002
 ; CURRENT APPLICATION NUMBER: US/09/817,774
 ; CURRENT FILING DATE: 2001-03-26
 ; PRIOR APPLICATION NUMBER: 60/192,202
 ; PRIOR FILING DATE: 2000-03-27
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 37
 ; LENGTH: 399
 ; TYPE: PRF
 ; ORGANISM: Arabidopsis thaliana
 US-09-817-774-37

Query Match 84.1%; Score 2182; DB 10; Length 399;
 Best Local Similarity 100.0%; Pred. No. 2e-197;
 Matches 399; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAETVSPITVYASMSLAFCEPFFVILLMTVHODGSVTQTFGEFWENGVOGLINIMP 60
 DB 1 MAETVSPITVYASMSLAFCEPFFVILLMTVHODGSVTQTFGEFWENGVOGLINIMP 60
 QY 61 RPTLIAMKIIIFCYGAFEAIIQLLPGKRVESPISPAGNRPVYKANGIAEFTLATYGL 120
 DB 61 RPTLIAMKIIIFCYGAFEAIIQLLPGKRVESPISPAGNRPVYKANGIAEFTLATYGL 120
 QY 121 WMGFIENPAIYDHLGELIFSGFIICVLYIKGHVAPSSDSCGNLIIDFYWGM 180
 DB 121 WMGFIENPAIYDHLGELIFSGFIICVLYIKGHVAPSSDSCGNLIIDFYWGM 180
 QY 181 ELYPRIGKSPDIKFTNCRFGMSMAVLAATYCIKOYEINGKVSMSLVNTIIMLVYTK 240
 DB 181 ELYPRIGKSPDIKFTNCRFGMSMAVLAATYCIKOYEINGKVSMSLVNTIIMLVYTK 240
 QY 241 FFWMEAGYNTMTDIAHDRGFEYICWGLVWPVSYSYTPGMVLVNHPELGTQLAITYI 300
 DB 241 FFWMEAGYNTMTDIAHDRGFEYICWGLVWPVSYSYTPGMVLVNHPELGTQLAITYI 300
 QY 301 GILCIYINDCDROQROEFRTNCKCLVWGRAPSKIVASTTTSGETKTSLLTSGWGM 360
 DB 301 GILCIYINDCDROQROEFRTNCKCLVWGRAPSKIVASTTTSGETKTSLLTSGWGM 360
 QY 361 RHFHVPEILSAFWTVPALFQNFLEYVIFLTLLEPDRAKDDDC 408
 DB 361 RHFHVPEILSAFWTVPALFQNFLEYVIFLTLLEPDRAKDDDC 408

RESULT 8

US-09-817-774-33
 ; Sequence 33, Application US/09817774
 ; Patent No. US2002012011A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHOE, Sunghwa
 ; APPLICANT: FELDMANN A., Kenneth
 ; TITLE OF INVENTION: Dwf5 MUTANTS
 ; FILE REFERENCE: 2225-0020 / 91020.002
 ; CURRENT APPLICATION NUMBER: US/09/817,774
 ; CURRENT FILING DATE: 2001-03-26
 ; PRIOR APPLICATION NUMBER: 60/192,202
 ; PRIOR FILING DATE: 2000-03-27

NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 33
 ; LENGTH: 298
 ; TYPE: PRF
 ; ORGANISM: Arabidopsis thaliana
 US-09-817-774-33

Query Match 61.9%; Score 1606; DB 10; Length 298;
 Best Local Similarity 100.0%; Pred. No. 1.9e-143;
 Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAETVSPITVYASMSLAFCEPFFVILLMTVHODGSVTQTFGEFWENGVOGLINIMP 60
 DB 1 MAETVSPITVYASMSLAFCEPFFVILLMTVHODGSVTQTFGEFWENGVOGLINIMP 60
 QY 61 RPTLIAMKIIIFCYGAFEAIIQLLPGKRVESPISPAGNRPVYKANGIAEFTLATYGL 120
 DB 61 RPTLIAMKIIIFCYGAFEAIIQLLPGKRVESPISPAGNRPVYKANGIAEFTLATYGL 120
 QY 121 WMGFIENPAIYDHLGELIFSGFIICVLYIKGHVAPSSDSCGNLIIDFYWGM 180
 DB 121 WMGFIENPAIYDHLGELIFSGFIICVLYIKGHVAPSSDSCGNLIIDFYWGM 180
 QY 181 ELYPRIGKSPDIKFTNCRFGMSMAVLAATYCIKOYEINGKVSMSLVNTIIMLVYTK 240
 DB 181 ELYPRIGKSPDIKFTNCRFGMSMAVLAATYCIKOYEINGKVSMSLVNTIIMLVYTK 240
 QY 241 FFWMEAGYNTMTDIAHDRGFEYICWGLVWPVSYSYTPGMVLVNHPELGTQ 292
 DB 241 FFWMEAGYNTMTDIAHDRGFEYICWGLVWPVSYSYTPGMVLVNHPELGTQ 292

RESULT 9

US-09-817-774-35
 ; Sequence 35, Application US/09817774
 ; Patent No. US2002012011A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHOE, Sunghwa
 ; APPLICANT: FELDMANN A., Kenneth
 ; TITLE OF INVENTION: Dwf5 MUTANTS
 ; FILE REFERENCE: 2225-0020 / 91020.002
 ; CURRENT APPLICATION NUMBER: US/09/817,774
 ; CURRENT FILING DATE: 2001-03-26
 ; PRIOR APPLICATION NUMBER: 60/192,202
 ; PRIOR FILING DATE: 2000-03-27
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 35
 ; LENGTH: 323
 ; TYPE: PRF
 ; ORGANISM: Arabidopsis thaliana
 US-09-817-774-35

Query Match 61.9%; Score 1606; DB 10; Length 323;
 Best Local Similarity 100.0%; Pred. No. 2.1e-143;
 Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAETVSPITVYASMSLAFCEPFFVILLMTVHODGSVTQTFGEFWENGVOGLINIMP 60
 DB 1 MAETVSPITVYASMSLAFCEPFFVILLMTVHODGSVTQTFGEFWENGVOGLINIMP 60
 QY 61 RPTLIAMKIIIFCYGAFEAIIQLLPGKRVESPISPAGNRPVYKANGIAEFTLATYGL 120
 DB 61 RPTLIAMKIIIFCYGAFEAIIQLLPGKRVESPISPAGNRPVYKANGIAEFTLATYGL 120
 QY 121 WMGFIENPAIYDHLGELIFSGFIICVLYIKGHVAPSSDSCGNLIIDFYWGM 180
 DB 121 WMGFIENPAIYDHLGELIFSGFIICVLYIKGHVAPSSDSCGNLIIDFYWGM 180
 QY 181 ELYPRIGKSPDIKFTNCRFGMSMAVLAATYCIKOYEINGKVSMSLVNTIIMLVYTK 240
 DB 181 ELYPRIGKSPDIKFTNCRFGMSMAVLAATYCIKOYEINGKVSMSLVNTIIMLVYTK 240

Qy 241 FFWMEAGYNTMDIAHNRGFFYICWGLVWVSVYTSRGMVNHPELGTQ 292
 Db 241 FFWMEAGYNTMDIAHNRGFFYICWGLVWVSVYTSRGMVNHPELGTQ 292

RESULT 10

US-09-817-774-27
 ; Sequence 27, Application US/09817774
 ; Patent No. US20020120111A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHOE, Sunghwa
 ; APPLICANT: FELDMANN A., Kenneth
 ; TITLE OF INVENTION: DWFS MUTANTS
 ; FILE REFERENCE: 2225-0020 / 91020.002
 ; CURRENT APPLICATION NUMBER: US/09/817,774
 ; PRIOR FILING DATE: 2001-03-26
 ; PRIOR APPLICATION NUMBER: 60/192,202
 ; PRIOR FILING DATE: 2000-03-27
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 27
 ; LENGTH: 471
 ; TYPE: PRF
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: S7R-RAT
 US-09-817-774-27

Query Match 27.4%; Score 710.5; DB 10; Length 471;
 Best Local Similarity 36.6%; Pred. No. 5e-59;

Matches 157; Conservative 85; Mismatches 158; Indels 29; Gaps 9;

Qy 13 ASMSLAFPCPFVILMTVMVHODGSSVOTF--GFEMENVOGLINW---PRPTLAK 68
 Db 37 ASIFELLFAP--FIVYTFMACDQYSCSLAPALDIATGSHASLADIWAKTPVTAKAQ 94
 Qy 69 IIFCYGAFEAIDLLP-----GKRVGPISPAQNPVYKANGLAAYFV----113
 Db 95 LVALMVSFQYLKSWLPDFCHRELPGVYGVGOGAITPAGVYKANGLAAYFV----154
 Qy 114 -LATYGLMFMFGJNPAIVYDH-LGEIFSAIIFGSEFICVLLYIKGHVAPSS--DSGSG 170
 Db 155 FVNAVYLLS---FSPITIIDNMIPLMCANILG-YAVSTFAMIKGLYPTSAEDCKFTG 209
 Qy 171 NLIDFWMGELPRIGKSPDIFVTNCRGMSMAVLAATYCIKOYEINGKYSDSMLN 230
 Db 210 NFEYNTAMGIEFNPRIKGFDFELFNGRPGIYAMTLINLSFAKQOELGHTNSMLIV 269
 Qy 231 TILMLVYVTKFFWMEAGYNTMDIAHNRGFFYICWGLVWVSVYTSRGMVNHPELGT 290
 Db 270 NVLQAIYVLDFFWNETWYLTKTIDICHDFGWLPGWGDVWLPYLYTLQGLYLVHVVOLS 329
 Qy 291 TQALYIVLVAIGLCIYINVCDDROERFRRTNGKCLVWGRAPSKYIVASTTSGEKTSL 350
 Db 330 TPNALGILGLVGYIIFRTNHOQKDLFRTDRCRLGCKPAICSYTSAGLGHNHK 389
 Qy 351 LITSGMGLARHFRVPELISAFETVPALDFNLAIFYVYIFLLLPFRKDDRCRS 410
 Db 390 LTVSGFMGVARHFNVTGDLMSLAVYCLACGGCHLLPYFYITTYITLTHRCRLDEHRCAN 449
 Qy 411 KYGKYMKLY 419
 Db 450 KYGRDWEY 458

RESULT 11

US-09-729-674-142
 ; Sequence 142, Application US/09729674
 ; Patent No. US20010039335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jacobs, Kenneth
 ; APPLICANT: McCoy, John M.

APPLICANT: Lavalie, Edward R.
 APPLICANT: Collins-Racie, Lisa A.
 APPLICANT: Evans, Cheryl
 APPLICANT: Merberg, David
 APPLICANT: Treacy, Maurice
 APPLICANT: Agostino, Michael J.
 APPLICANT: Steinger II, Robert J.
 APPLICANT: Spaulding, Vikki
 APPLICANT: Wong, Gordon G.
 APPLICANT: Clark, Hilary
 APPLICANT: Fectel, Kim
 APPLICANT: Genetics Institute, Inc.
 ; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
 ; FILE REFERENCE: 6055-64X
 ; CURRENT APPLICATION NUMBER: US/09/729,674
 ; PRIOR FILING DATE: 2000-12-04
 ; PRIOR APPLICATION NUMBER: 09/539,330
 ; PRIOR FILING DATE: 2000-03-30
 ; NUMBER OF SEQ ID NOS: 283
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 142
 ; LENGTH: 475
 ; TYPE: PRF
 ; ORGANISM: Homo sapiens
 US-09-729-674-142

Query Match 26.9%; Score 698.5; DB 10; Length 475;
 Best Local Similarity 35.3%; Pred. No. 6.7e-58;
 Matches 154; Conservative 87; Mismatches 152; Indels 43; Gaps 10;

Qy 13 ASMSLAFPCPFVILMTVMVHOD-----GSYOTRFGFWENGVOGLINW---PRP 62
 Db 41 ASYIFELLFAP--FIVYTFMACDQYSCALTPVVDIV---TGARLSIDIMAKTPPI 92
 Qy 63 TILAKIIFCYGAFEAIDLLP-----GKRVGPISPAQNPVYKANGLAAYF 111
 Db 93 TRKAQLYLVMTFQVLYTSLPDCCHKFLPGYVGIGEGANTPAGVYKANGLAAYF 152
 Qy 112 VTLATYGLMFMFGJNPAIVYDH-LGEIFSAIIFGSEFICVLLYIKGHVAPSS 164
 Db 153 LF---HLMPFANMHLDSWSPSTIIFDNWIPILMCANILG-YAVSTFAMIKGLYPTSA 206
 Qy 165 -DSGSGNLIIDFWMGELPRIGKSPDIFVTNCRGMSMAVLAATYCIKOYEINGKYS 223
 Db 207 RCKPFGNFFYVYKMGIEFNPRIKGFDFELFNGRPGIYAMTLINLSFAKQOELGHTNS 266
 Qy 224 SDSMLVNTILMLVYVTKFFWMEAGYNTMDIAHNRGFFYICWGLVWVSVYTSRGM 283
 Db 267 TNANVLVNLQAIYVLDFFWNETWYLTKTIDICHDFGWLPGWGDVWLPYLYTLQGLY 326
 Qy 284 NHPVELGQALYIIVAGLCIYINVCDDROERFRRTNGKCLVWGRAPSKYIVASTTSG 343
 Db 327 YHPVOLSTPHAVGVLLGLVGYIIFRTNHOQKDLFRTDRCRLGCKPAICSYTSAGLGH 386
 Qy 344 GETKSLITSGMGLARHFRVPELISAFETVPALDFNLAIFYVYIFLLLPFRKDDRCRS 403
 Db 387 GORHNSKILLSGFMGVARHFNVTGDLMSLAVYCLACGGCHLLPYFYITTYITLTHRCRL 446
 Qy 404 DDDRCRSKYGYMKLY 419
 Db 447 DEHRCASKYGRDWEY 462

RESULT 12

US-09-817-774-26
 ; Sequence 26, Application US/09817774
 ; Patent No. US20020120111A1
 ; GENERAL INFORMATION:
 ; APPLICANT: CHOE, Sunghwa
 ; APPLICANT: FELDMANN A., Kenneth
 ; TITLE OF INVENTION: DWFS MUTANTS
 ; FILE REFERENCE: 2225-0020 / 91020.002
 ; CURRENT APPLICATION NUMBER: US/09/817,774

CURRENT FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: 60/192,202
 PRIOR FILING DATE: 2000-03-27
 NUMBER OF SEQ ID NOS: 45
 SOFTWARE: Patent Ver. 2.0
 SEQ ID NO 26
 LENGTH: 475
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: STR-HUMAN
 US-09-817-774-26

Query Match 26.9%; Score 698.5; DB 10; Length 475;
 Best Local Similarity 35.3%; Pred. No. 6,7e-58;
 Matches 154; Conservative 87; Mismatches 152; Indels 43; Gaps 10;

13 ASMLSLAFCPPEVILLMTWVHOD-----GSVTFQGFENGVOGLINIV---PRP 62
 41 ASVIFLLLPAP---FIVYFIMACDQYSCALTPGVVDIV---TGHARLSIDIMAKTPPI 92
 63 TLIAKIIIFCYGAFEAIIQDLLP-----GKRVGPISPAGNRPYKANGLAAYF 111
 93 TRKAAGLYTLWTFQVLLTSLPDCHEKFLPGYVGIGIQGAVTPAGVYVANKQINGLOAML 152
 112 VTLATYGLMWF-----GIFPAIYVDH-LGEIFSALIFGFCVLLYIKGHVAPSSS 164
 153 LT-----HLMTFANAHLSWFSPTIIFDWMIPILMCANILG-YAISTFAMVGVFFPISA 206
 165 -DSGSGNLIDFYWGMELYPRIKSEDIKFTNCRFGMSNAVLAVYCIKQYINGKV 223
 207 RCKEFTGNFYYNMGIEFNPRIGKWFDEKLFENGPRGIVAWTLINLSFAAKORELHSHV 266
 224 SOSMLVNTLIMLYVYTKFEFMEAGYWNMTMDIAHDSGEYICGCLVWPSVYTSQMYLY 283
 267 TNAMVAVNLOAIYVIDEFWNETWYTKTIDICHDFGMLGDCVWLPYLTLOGILYV 326
 284 NHEVELGTOLAIYILVAGLICIYINDCROREFRRTNGKCLVWGRAPSKIVASYTTTS 343
 327 YHVQVSTPHAVGVLLGLVGYIIFRVANHQDLPFRDTGRCILIMRKRVIECSYTSAD 386
 344 GEFKTSLLTSGWGLARHFIHYPELLSAFMTVPALFENFLAIFYVLTLLLEDRACK 403
 387 GORHNSKLLVSGFWYARHFNHYDLMGSLAYCLACGGHLLPYEIIYMAILTHRCIL 446
 404 DDDRCRKYKGYKLY 419
 447 DEHRCAKSKYGRMERY 462

RESULT 13
 US-09-817-774-23
 Sequence 23, Application US/09817774
 Patent No. US2002012011A1
 GENERAL INFORMATION:
 APPLICANT: CHOE, SungHwa
 APPLICANT: FELDMANN A., Kenneth
 TITLE OF INVENTION: DWFS MUTANTS
 FILE REFERENCE: 2225-0020 / 91020.002
 CURRENT APPLICATION NUMBER: US/09/817,774
 CURRENT FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: 60/192,202
 PRIOR FILING DATE: 2000-03-27
 NUMBER OF SEQ ID NOS: 45
 SOFTWARE: Patent Ver. 2.0
 SEQ ID NO 23
 LENGTH: 427
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: LBR-RAT
 US-09-817-774-23

Query Match 21.3%; Score 553.5; DB 10; Length 427;
 Best Local Similarity 33.0%; Pred. No. 2,5e-44;
 Matches 138; Conservative 65; Mismatches 176; Indels 39; Gaps 11;

16 LSLIAFCPPVILLMTWVHODGSVYQ-----TFQEFWNGVOGLINIMPRPTLLAMKI 69
 28 LLIMLGIPACVFLLLQCAQKDFELQFPPLPALRELMKAVRQVYLLM-----77
 70 IFYGAFAEIIQLLPGKRVGPISPAGNRPYKANGLAAYFVTLAVYGLMFGIFNPA 129
 78 FE-----LQALFSLIPYKRYVEGPIVDGRRLKRLNQLVAFILTSAA-VGTAFFMIDELY 132
 130 IVDHGEIFSALIFGFCVLLYIKGHVAPSSSDS-GSGCMLIDFYWGMELYPRIK 188
 133 YLYTHFIQFALAAIVFVSVLVSILYARSLKVPREDLSPASSGAAVYDFIGRELNPRIG- 191
 189 SFDIKVTNCRFGMSNAVLAVYCIKQYINGKVSQSM---LVNITLMVYVTKFEFWE 245
 192 AFDLKFCELRPGILGVVAVINLVNLMKRVQERSAPSLAMTLVNS-FOLLVYVDALME 250
 246 AGYWNMTDIAHDSGEYICGCLVWPSVYTSQMYLVNHPVELGTOLAIYILVAGLIC 305
 251 EALLTMDIHDGFMILAFGLDVTWVPTYSIQAFYLVNHPQLSMPLT-SVITALKLG 309
 306 YINYDC-DROREFRRTNGKCLVWGRAPSKIVASYTTTSGEYTSLLISGWMGLARH 364
 310 YVIFRCAVNSQKNAFRKN-----PTDKLAHLTKTIPSTWKSLLVSGWGFVHRPN 359
 365 YVPEILSAFMTVPALFENFLAIFYVLTLLLEDRACKRDDCRSKYKGYKLYCEK 422
 360 YLGLIMALAWSLPCGNHLLIPYVLYFTALLIHRARDEHCRKRYGLAMKRYCOR 417

RESULT 14
 US-09-817-774-22
 Sequence 22, Application US/09817774
 Patent No. US2002012011A1
 GENERAL INFORMATION:
 APPLICANT: CHOE, SungHwa
 APPLICANT: FELDMANN A., Kenneth
 TITLE OF INVENTION: DWFS MUTANTS
 FILE REFERENCE: 2225-0020 / 91020.002
 CURRENT APPLICATION NUMBER: US/09/817,774
 CURRENT FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: 60/192,202
 PRIOR FILING DATE: 2000-03-27
 NUMBER OF SEQ ID NOS: 45
 SOFTWARE: Patent Ver. 2.0
 SEQ ID NO 22
 LENGTH: 426
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: LBR-HUMAN
 US-09-817-774-22

Query Match 20.8%; Score 540; DB 10; Length 426;
 Best Local Similarity 32.7%; Pred. No. 4,6e-43;
 Matches 142; Conservative 65; Mismatches 157; Indels 70; Gaps 16;

8 PIIVYASML-----SLIACPPVILL--LMTWVHODGSVYQTFQEFWNGVOGLINI 58
 34 PVFLFLMLCMKQKDPSSLNPPPALPYELMETRY-----FQVY-----73
 59 WPRPTLLAMKIIFCYGAFEAIIQDLLP-GKRVGPISPAGNRPYKANGLAAYFVTLATY 117
 74 -----LIMFLI-----QVLFILPPIGKYVEGTPILDRRLKRLNQLVAFILTSAAV- 119
 118 LGLMWFGLFNPDAIYVDHGEIFSALIFGFCVLY-----LYIKGHVAPSSSDS-GSGCML 172
 120 IGTSLFGQVEFHYVYSH-----FLOFALAAVFCVLLSVYILMRSILKAPRNDLSPASSGNA 175
 173 IIDFYWGMELYPRIKSEDIKFTNCRFGMSNAVLAVYCIKQYINGKVSQSM---LV 229

Db 176 YDDEFIGRELNPRIG-TEFLAYFCELRPLIGWVWINLMLEAEKIDDRAPVSIAMLV 234
QY 230 NTILMLVVTKEFWMEAGYMTMDIAHNRGGEYICMGVLVWPVSYYTSPGYLVNHPVEL 289
Db 235 NS-FOLLIVVDALMNEELTLTMDIHDGFGMLAFGLVWPVPLYSFOAYLVSHPMVEV 293
QY 290 GTOLAIYILVAGILCIYIN-DCDROROEFRRTNGKCLWGRAPSKIYASYTTSGETKT 348
Db 294 SMPMASLIIVLK-LCGYIFRGANSOKNAFRKN-----PSDKLAHLKTIHTSSG 342
QY 349 SLLTSGWGLARHGHVPELISAFWTPALFDNFAVYVIFLTLFDRARDDRC 408
Db 343 KNLVSGWGWGFVRHBNYIGDLIMALMSLPCGFNHILPEFYIYETMLVHREARDEXHC 402
QY 409 RSKYGYKWLCEK 422
Db 403 KKGIVAMEKYCOR 416

RESULT 15
US-09-817-774-25
; Sequence 25 Application US/09817774
; Patent No. US2002012011A1
; GENERAL INFORMATION:
; APPLICANT: CHOE, Sunghwa
; APPLICANT: FELDMANN A., Kenneth
; TITLE OF INVENTION: Dwf5 MUTANTS
; FILE REFERENCE: 2225-0020 / 91020.002
; CURRENT APPLICATION NUMBER: US/09/817,774
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: 60/192,202
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 424
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: S14R-YEAST
US-09-817-774-25

Query Match 19.3%; Score 500.5; DB 10; Length 424;
Best Local Similarity 32.5%; Pred. No. 2,4e-39;
Matches 128; Conservative 67; Mismatches 172; Indels 27; Gaps 11;

QY 43 TFGFFWENGVOG-----LINWPRPLIAWKIIFCYGAFEALLOLL--LPGRVYEG- 91
Db 32 SFGSFYICNEBGCAPKFSKISHIKTPLFDOKSLIYLLMFSTLLMKCTNGKMAKGT 91
QY 92 PISPAQNRPPYKAGL-AAAYVTLATYLGMLWGFENPAIYDH-LGEISALIFGSEIF 149
Db 92 PIDDKGTRLRKINGFNSACILIGVCTSIYLGASCMERIMDNFIQLMFAAYVF-SVL 150
QY 150 CVLLYINGHVA PSS-SDSGSCNLIIDPYWGMELYPRIKSPDIKVTNCRFGMSNAVL 208
Db 151 CTFCYVOSFEKGQOLAKGTSNLLFDMFGRSLNPRIG-NFDIKCFELRPGILIMVVF 209
QY 209 AVTYCICQY-EINGKVSDSLVTIIMLVYVTFEWMWAGYMTMDIAHNRGGEYICWGC 267
Db 210 DIAFAHQYVLGRIIDSMVLIIFHTWYVLDLINESAVLITMDITTDGFGYMLSGFD 269
QY 268 LWWPVSYYTSPGMYLVNHPVELGTOLAIYILVAGILCIYINYDODROROEFRRTNGKCLV 327
Db 270 LWWPFLYSLQARYLAHFHVDLGLVKTALICLOFLGYIFRGANGOKNFRSPND-- 326
QY 328 WGRAPSKIYASYTTSGETKTSLTSGWGLARHGHVPELISAFWTPALFDNPLAY 387
Db 327 ----PKLHLKFIOTKRGTK---LITSGWGMAMHINYPGDWIMAMACLPAGFGSPIFY 379
QY 388 FYVIFLTLLEFDRAKRDDRCRSKYGYWKLYCE 421

Db 380 FYVAYFVGLLVHNRARDHRCRVKYGEDMEKYCK 413

Search completed: January 14, 2003, 14:21:34
Job time: 15 secs

